

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: Puliluker Type of Search 1 Vendors and cost where applicable
Searcher Phone #: 308-4740 NA Sequence (#) 2 STN _____
Searcher Location: 1507 AA Sequence (#) 2 Dialog _____
Date Searcher Picked Up: 6-29 Structure (#) _____ Qucstel/Orbit _____
Date Completed: 7-2-01 Bibliographic _____ Dr. Link _____
Searcher Prep & Review Time: 5 Litigation _____ Lexis/Nexis _____
Clerical Prep Time: _____ Fulltext _____ Sequence Systems 01, 02
Online Time: 5 Patent Family _____ WWW/Internet _____
Other _____ Other (specify) _____

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2001, 23:49:58 ; Search time 1372.68 Seconds
(without alignments)
12147.648 Million cell updates/sec

Title: US-09-457-066-1
Perfect score: 1764
Sequence: 1 attatgtggaactacccctg.....cttgtgtcgtgtagtagga 1764

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

| Database : | EST : |
|------------------|-------------------|
| 1: gb_est1.* | 44: em_esthum10.* |
| 2: gb_est2.* | 45: em_esthum11.* |
| 3: gb_est3.* | 46: em_esthum12.* |
| 4: gb_est4.* | 47: em_esthum13.* |
| 5: gb_est5.* | 48: em_esthum14.* |
| 6: gb_est6.* | 49: em_esthum15.* |
| 7: gb_est7.* | 50: em_esthum16.* |
| 8: gb_est8.* | 51: em_esthum17.* |
| 9: gb_est9.* | 52: em_esthum18.* |
| 10: gb_est10.* | 53: em_esthum19.* |
| 11: gb_est11.* | 54: em_esthum20.* |
| 12: gb_est12.* | 55: em_esthum21.* |
| 13: gb_est13.* | 56: em_esthum22.* |
| 14: gb_est14.* | 57: em_esthum23.* |
| 15: gb_est15.* | 58: em_esthum24.* |
| 16: gb_est16.* | 59: em_esthum25.* |
| 17: gb_est17.* | 60: em_esthum26.* |
| 18: gb_est18.* | 61: em_esthum27.* |
| 19: gb_est19.* | 62: em_esthum28.* |
| 20: gb_est20.* | 63: em_estin1.* |
| 21: gb_est21.* | 64: em_estin2.* |
| 22: gb_est22.* | 65: em_estin3.* |
| 23: gb_est23.* | 66: em_estin4.* |
| 24: gb_est24.* | 67: em_estin5.* |
| 25: gb_est25.* | 68: em_estom1.* |
| 26: gb_est26.* | 69: em_estom2.* |
| 27: gb_est27.* | 70: em_estov1.* |
| 28: gb_est28.* | 71: em_estov2.* |
| 29: gb_est29.* | 72: em_estpl1.* |
| 30: gb_est30.* | 73: em_estpl2.* |
| 31: gb_est31.* | 74: em_estpl3.* |
| 32: gb_est32.* | 75: em_estpl4.* |
| 33: em_estba.* | 76: em_estpl5.* |
| 34: em_estfun.* | 77: em_estpl6.* |
| 35: em_esthum1.* | 78: em_estpl7.* |
| 36: em_esthum2.* | 79: em_estpl8.* |
| 37: em_esthum3.* | 80: em_estpl9.* |
| 38: em_esthum4.* | 81: em_estpl10.* |
| 39: em_esthum5.* | 82: em_estro1.* |
| 40: em_esthum6.* | 83: em_estro2.* |
| 41: em_esthum7.* | 84: em_estro3.* |
| 42: em_esthum8.* | 85: em_estro4.* |
| 43: em_esthum9.* | 86: em_estro5.* |
| | 87: em_estro6.* |
| | 88: em_estro7.* |
| | 89: em_estro8.* |
| | 90: em_estro9.* |
| | 91: em_estro10.* |
| | 92: em_estro11.* |
| | 93: em_estro12.* |
| | 94: em_estro13.* |
| | 95: em_estro14.* |
| | 96: em_estro15.* |
| | 97: em_estro16.* |
| | 98: em_estro17.* |
| | 99: em_estro18.* |
| | 100: em_estro19.* |
| | 101: em_estro20.* |
| | 102: gb_est25.* |
| | 103: gb_est26.* |
| | 104: gb_est27.* |
| | 105: gb_est28.* |
| | 106: gb_est29.* |
| | 107: gb_est30.* |
| | 108: gb_est31.* |
| | 109: gb_est32.* |
| | 110: gb_est41.* |
| | 111: gb_est42.* |
| | 112: gb_est43.* |
| | 113: gb_est44.* |
| | 114: gb_est45.* |
| | 115: gb_est46.* |
| | 116: gb_est47.* |

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.


```

Db 533 TTCTAGTGTACTACAAACCTGGTTTAAAGAGGAACTATGTGCTATGAACCTAAACT 592
Qy 1747 tgtgtctgtctgatagga 1764
|||||
Db 593 GGTCTGCTGTGATAGGA 610
|||||

RESULT 2
BE326924/c 500 bp mRNA EST 14-JUL-2000
LOCUS hr-67f08.x1 NCI_CGAP_kid11 Homo sapiens cDNA clone IMAGE:3133575 3',
DEFINITION mRNA sequence.
ACCESSION BE326924
VERSION BE326924.1 GI:9200700
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 500)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 462.
FEATURES
source
1. 500
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3133575"
/clone_lib="NCI_CGAP_kid11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 132376-132391, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 163 a 98 c 90 g 149 t
ORIGIN
Query Match 27.6%; Score 486.4; DB 166; Length 500;
Best Local Similarity 99.4%; Pred. No. 4.6e-122;
Matches 499; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Qy 1258 aacgtatcgttatctccatcttaatctcagttgttgcctcaggacaccttcattcttc 1317
|||||
Db 500 AACGTATCGGTATCTCCATCTTAATCTCAGTTGTTGCTTCAAGGACCTTCATCTTC 441
|||||

Qy 1318 aggtattacagtgatctctgaagagagagacatcaacagaaattaggtgtgtcaacag 1377
|||||
Db 440 AGGATTTACAGTGCTATCTCTGAAGAGGATACATCAACACAGAAATAGGAGTTGTGCAACAG 381
|||||

Qy 1378 ctcttttgagaggagcctaaaggacaggaagaaaggtcttcaatctgtggaagaaatt 1437
|||||
Db 380 CTCCTTTTGAGAGGAGGCTTAAAGACAGAGGAAAGAGTCTCAATCTGGAAGAAAGAAATT 321
|||||

```

```

Qy 1438 aaatgtttattaaatagatcaccagctagtttcagagttaccatgtacgtattccacta 1497
|||||
Db 320 AAATGTTTATTAATAGATACACAGCTAGTTTCAGAGTTACCATGTACGTATTCACACTA 261
|||||

Qy 1498 gctgggttctgtatttcagttcttcgatacggcttaggtaagtgtcagtagcaggaataa 1557
|||||
Db 260 GCTGGGTTCTGTATTTTCAGTTCTTCGATACGGCTTAGGCTAATGTCATGACAGGAAAAA 201
|||||

Qy 1558 aactgtgcaagtgcacactgattccgttgcttgccttaaacctcaaaagctccatgtcctg 1617
|||||
Db 200 AACTGTGCAAGTGAGCAGCAGCTGATTCGTTGCGCTTAACCTCTAAAGCTCCATGTCCTG 141
|||||

Qy 1618 ggcctaaatcgatataaaatctggatttttttttttttttttttttttttttttttttttt 1677
|||||
Db 140 GGCCATAAATCGTATAAAATCTGGA--TTTTTTTTTTTTTTTTTTTGTCTCATATTACATATG 83
|||||

Qy 1678 taaccagagaacattctatgtactacaacactggtttttaaagaagaaactatgtgtgtatg 1737
|||||
Db 82 TAACCAGAACATCTCTATGTACTACAACCTGGTTTTTAAAAAGGAACTATGTGTGCTATG 23
|||||

Qy 1738 aattaaactgtgtcgtgctga 1759
|||||
Db 22 AATTAAACTTGTGTCATGCTGA 1
|||||

RESULT 3
BE958470 523 bp mRNA EST 04-OCT-2000
LOCUS 601644787F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:3930045 5',
DEFINITION mRNA sequence.
ACCESSION BE958470
VERSION BE958470.1 GI:10569175
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 523)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM763 row: m column: 22
High quality sequence start: 2
High quality sequence stop: 513.
FEATURES
source
1. 523
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3930045"
/clone_lib="NIH_MGC_56"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccattagggc); Site_2: SfiI (ggccattagggc);
Double-stranded cDNA was prepared from cell line RNA.
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGGCGGCGGCGGACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 147 a 118 c 129 g 128 t 1 others

```

ORIGIN

Query Match 27.3%; Score 482.2; DB 142; Length 523;
Best Local Similarity 98.7%; Pred. No. 6.6e-121;
Matches 517; Conservative 0; Mismatches 4; Indels 3; Gaps 3;
QY 677 tgcacccttcagcttgccactggaactgcttaataatgctataactgaccttagta 736
Db 1 TGTACCCCTTCAGC-TTGCCACTGGACTGCTTAATAATGCTATAACTGCTTTAGTA 59
QY 737 ccttggaagacattatccatcttgcagccagagagatggcagtggtgacttagaagtc 796
Db 60 CCTTGGAGACCTTATTCGATATCTTGAACACAGAGATGGCAG-TCGACTTGAAGATC 118
QY 797 tatatagccaaactggcaactcttggcaagctttgtttttggaagaaataccagag 856
Db 119 TATATAGGCCAACTTGGCAACTCTTGGCAAGCTTCTGTTTTCGGAAGAAATCCAGAG 178
QY 857 tggtagctgaacctcttaacagagaggtgaagattatatacagctcacacctcgtaact 916
Db 179 TGGTGGATCTGAACCTTCTTAACAGAGGAGGTAGATATATACAGCTGCACACCTTCGTAAC 238
QY 917 tctcagtgctccaaaggggaactaaagaaagacccgataccattttctggccaggtgtc 976
Db 239 TCTCAGTGTCCTAAGGGAAGAACTAAACAGAACCCGATACCATTTTCTGCCAGGTGTC 298
QY 977 tctcagtgtaaacgctgtggtggaactgtgcctgtgtgtctccacaattgcaatgaatgc 1036
Db 299 TCTCGTAAACGCTGTGGTGGAACTGTGCCTGNTGTCTCCCAATTCGAATGAATGTC 358
QY 1037 aatgtgtcccaagcaaaacttaaaaaataccagagctcttcagttgagaccagaaga 1096
Db 359 AATGTGTCCCAAGCAAGTACTAAAAATACCAAGAGCTCTTCAGTTGAGACCAAGA 418
QY 1097 ccggtgtcagggggtgcacaaactcactcaccgagctggccctggagccactgagaggt 1156
Db 419 CCGGTGTACAGGGATTGACAAATCACTCACCGACGTGG-CCTGGAGACCATGAGAGT 477
QY 1157 gtgactgtgtgcagagggagacagagggatagccgcatacac 1200
Db 478 GTGACTGTGTGTCAGAGGGAGCAGAGGAGTAGCCGAATCAC 521

RESULT 4

BF102859 851 bp mRNA EST 19-OCT-2000
LOCUS 60164827F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4073095 5',
DEFINITION mRNA sequence.

ACCESSION BF102859
VERSION BF102859.1 GI:10885385

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 851)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: DCFD/DPF

cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L10M921 row: f column: 08

High quality sequence stop: 529.

Location/Qualifiers

1..851

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4073095"
/clone_lib="NIH_MGC_60"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggccattatgcc)
Site_2: SfiI (ggccattatgcc)
); Double-stranded cDNA was prepared from cell line RNA.
5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCCGCGGCACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT 234 a 169 c 238 g 210 t
ORIGIN

Query Match 27.1%; Score 478.8; DB 144; Length 851;
Best Local Similarity 96.6%; Pred. No. 6.3e-120;
Matches 543; Conservative 0; Mismatches 12; Indels 7; Gaps 5;
QY 639 tgcattgccacaattcacagaagctgtgagctcttcagtgctaccccttcagctttgcc 698
Db 8 TGTATGCCACAATTCAGAAAGCTGTGAGTCTCTTCACTGTACCCCTTTCAGC-TTGGC 66
QY 699 actggacctcttaataatgctataactgcttttagtaccttggaagaccttattcgata 758
Db 67 ACTGGACCTGCTTAATAATGCTATAACTGCTTTAGTACCTTTGGAAGACCTTATTTCGATA 126
QY 759 tcttgaccagagagatggcagttggacttagaagctatataagccaaacttggcaact 818
Db 127 TCTTGAACACAGAGAGATGGCAGTTGGACTTAGAAGATCTATATAGCCAACTTGGCAACT 186
QY 819 tcttgcaagagctttgtttttggaagaaatccagagtggtgagctctgaaaccttctaac 878
Db 187 TCTTGGCAAGGCTTT--GTTTGGAAAGAAATCCAGAGTGGTGGATCTGAACCTTCTTAC 244
QY 879 agagaggtgaagattatatacagctcacacctcgtaaccttcagtgctcataagagaaga 938
Db 245 AGAGAGGTGAAGATTATACAGCTGCACACCTCGTAACCTTCAGTGTCTCAAGGGAAGA 304
QY 939 actaaagagaaccgataccattttctggccaggtgtgtctcctgggttaaacgctgtggtg 998
Db 305 ACTAAAGAGAACCGATACCATTTTCTGCCAGGTGTCTCTCTGGTTAAACGCTGTGGTGG 364
QY 999 gaactgtcctgtgtctccacaattgcaatgaatgtcaatgtgtcccaagcaaggttac 1058
Db 365 GAACGTGTGCTGTG-TGTCTCCACAATTCGAATGAATGTCTCAATGTCTCCCAAGCAAGTTAC 423
QY 1059 taaaaataccagagctcttcagttgagaccaaagaccggtgtcagggattacaa 1118
Db 424 TAAAAATACCAACAGGTCCTTCAGTTGAGACCAAGACCGGTGTCTCAGGGGA-TGCACAA 482
QY 1119 atcaactcaccgagctggccctggagcaccatgagagtggtgactgtgtgagagggag 1178
Db 483 ATCACTACCAGCTGGGCC--TGAGCACCATGAGAGGTGTGACTGTGTCTGCGAGGAGC 540
QY 1179 cacagaggatagccgcatacac 1200
Db 541 CACAGGGGATAGCCGATTAAAC 562

RESULT 5

BE884591

LOCUS

DEFINITION

ACCESSION

BE884591 565 bp mRNA EST 20-OCT-2000
601506290F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3907918 5',
mRNA sequence.
BE884591

[illegible]

ACCESSION BF697294
VERSION BF697294.1 GI:11982702
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 820)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC alone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
Plate: LCM1123 row: o column: 08
High quality sequence stop: 462.
FEATURES
Source
1..820
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4286815"
/clone_lib="NIH MGC_56"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pBR-LIB (Clontech); Site: 1:
SfiI (ggccgcttggcc); Site: 2: SfiI (ggccattatggcc);
Double-stranded cDNA was prepared from cell line RNA.
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGCCATTATGGC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCCGAGGCGCGCATG-df(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 221 a 179 c 195 g 225 t
Query Match 25.8%; Score 455.6; DB 168; Length 820;
Best Local Similarity 93.6%; Pred. No. 1.4e-113;
Matches 498; Conservative 1; Mismatches 25; Indels 8; Gaps 2;
QY 1235 ygcagtggctgatttattagaaacgtatgcgtttatcttccatccttaattcagttgtt 1294
:|||||
Db 13 TGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCTCTTAATCTCAGTTGTT 72
QY 1295 tgcctcaaggaccttcattccttcagagattacagtgcatcttgaaaggagacatcaaa 1354
:|||||
Db 73 TGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATCTCTGAAAGAGGAGACATCAA 132
QY 1355 cagaattaggagttgtgcaacagctctttgagagagagccttaaaaggacagagaaaag 1414
:|||||
Db 133 CAGAAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGAGGCGCTTAAAGGACAGAGAAAGG 192
QY 1415 tottcaatcgtggaagaaaaataaattgtgtattataatagatcacaccagctagtttcaga 1474
:|||||
Db 193 TCTTCAATCGTGGAAAGAAAATTAATGTGTATTAAATAGATCACCAGCTAGTTTCAGA 252
QY 1475 gttaccatgacgtattccactagctggttctgtatttcagttcttttcagtcacgctta 1534
:|||||
Db 253 GTTACCATGTAGTATTCCACTAGCTAGTGGGTCTGTATTTCAGTTCTTTTCGATACGCTTA 312
QY 1535 gggttaattcagtcagaaaaaactgtgaagtgcagacctgattccgttgccttgcct 1594
:|||||
Db 313 GGGTAATGTACGACAGGAAAAACCTGTGCAAGTGAGCACCTGATTCGTTGCTTGCT 372

QY 1595 taactctaaagctccatgtctctgggcctaaatcgtataaaatcgtgattttttttt 1654
:|||||
Db 373 TAACTCTAAAGCTCCATGTCTCTGGCGCTAAATCGTATAAAATCTGGA-----TTTTT 426
QY 1655 tttttgctcatattcacatattgtaaacaccagaacattctatgtactacaaacacctggttt 1714
:|||||
Db 427 CTTTGGGCCCATTTTCCTTTTGTAAACCCAGAACTTTCTATGCTCTTAACAACCTGGTTT 486
QY 1715 taaaaaggaa--ctatgttctgatgaattaaacctgtgtgtgtgtgtataga 1764
:|||||
Db 487 TCAACACGGACCCCTATGTGCTGATGAATTAACCTGGTGTGCTGTATAGGA 538
RESULT 9
BG609411 556 bp mRNA EST 17-APR-2001
LOCUS 323251 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
DEFINITION BG609411
ACCESSION BG609411
VERSION BG609411.1 GI:13659390
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 556)
REFERENCE 1
AUTHORS Fahrrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keeler,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt.trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR PRIMERS
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCCAGTCACGACG
Plate: 98 row: P column: 8
Seq primer: ATTTAGTGACACTATAG.
FEATURES
Source
1..556
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1PIG"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site: 1: XbaI; Site: 2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT 139 a 153 c 141 g 123 t
ORIGIN
Query Match 25.2%; Score 444; DB 155; Length 556;
Best Local Similarity 87.4%; Pred. No. 1.9e-110;
Matches 486; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 615 agggttctgcattccattcaaacattgtcattgccacattcacagaagcttgaccttc 674
:|||||
Db 1 AGGTTCTGCATCCATACAAATCGTCACACCAATTCACAGAACTGTGAGTCCCTC 60
QY 675 agtgcacccttcagcttggcactggacctgctttaaataatgataactgccttag 734
:|||||
Db 61 ACTGCTACCGCTGCAGCTCTGCCCTGGACCTGCTTAACAATGCTGCTACTGCCTTAG 120
QY 735 taccctggagaccttattogatattttgaaccagagagatggcagttggacttagaaga 794
:|||||

Db 121 AGATTGTATCGATGAATATTTCCCTTCTAACCCAGGGTTCTGCATCCACTTACACAT 180
Qy 639 tgtcatgccacaattcacagaagctgtgagctctcagtgctaccccccttcagctttgcc 698
Db 181 TGTCTATGCCCAATTCACAGAAGCTGTGAGTCTCTCAGTGTCTACCCCTTCAGCTTTGCC 240
Qy 699 actgacctgcttaataatgcataactgcctttagctaccttgggaagaccttattcgata 758
Db 241 ACTGGACCTGCTTAATAATGCTATAACTGCCCTTAGTACCTTGGAGAGACTTATTTCGATA 300
Qy 759 tcttgaaccagagagatgacagcttgagcttagaagatctatatagggccaacttggcaact 818
Db 301 TCTTGAACCAGAGATGCGACCTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACT 360
Qy 819 tcttgcaagcttctgttttttggagaataatccagagtggtggtatctgaa 869
Db 361 TCTTGGCAAGGCTTTGTGTTTGGAGAAATCCAGAGTGGTGGATCTGAA 411

RESULT 13
BF151355 564 bp mRNA EST 29-DEC-2000
LOCUS uz15b12.y1 NCI-CGAP_Mam5 Mus musculus cDNA clone IMAGE:3669119 5'
DEFINITION similar to FR:Q90Y71 Q90Y71 FALLOTEIN. ; mRNA sequence.
ACCESSION BF151355
VERSION BF151355.1 GI:11032750
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 564)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

MGI:1429887
Seq primer: -40RP from Gibco
High quality sequence stop: 436.
Location/Qualifiers
1. 564
/organism="Mus musculus"
/strain="C57/B6"
/db_xref="taxon:10090"
/clone="IMAGE:3669119"
/clone_lib="NCI-CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

BASE COUNT 157 a 130 c 140 g 137 t
ORIGIN

Query Match 23.1%; Score 407.8; DB 145; Length 564;
Best Local Similarity 82.8%; Pred. No. 1.5e-100;
Matches 466; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

Qy 619 tctgtccactacacaattgttcacacaaatttcacagaagctgtgagctcttcagtg 678
Db 1 TTTGTGCATCCACTACAGTATTATTCATGCCACAAGTCACAGAAACACACGAGTCTTCGCTG 60
Qy 679 ctaccccttcagcttggcaactgacctgcttaataagctataaactgcttagtacc 738
Db 61 TTGCCCCCTTTCATCTTTTGTCTATTGGACCTGCTCAACATGCTGACTGCTTTCAGTACC 120
Qy 739 ttggaagaccttattcgatatcttgaaccagagagatggcagcttgaccttagagatata 798
Db 121 TTGGAAGAGCTGATTCGGTACCTAGAGCCAGATCGATGGCAGGTGGACTTTGGACAGCCTC 180
Qy 799 tataggccaacttggcaactcttggcaagcttttgggttttttggagaagaaatcccagatg 858
Db 181 TACAAGCCACATGGCAGCTTTTGGCAAGCTTTTCTGTATGGAAAAAAGCAAAATG 240
Qy 859 gtggatctgaaccttctacagagaggttaagattatatacagctgcacaccttgtaacttc 918
Db 241 GTGAATCTGAATCTCTCAAGGAAGAGGTAAACTCTACAGCTGCACACCCCGGAACCTC 300
Qy 919 tcaagtccataaagggaagaactaaagaaacagacacacattttctgcccaggttgcttc 978
Db 301 TCAGTGTCATACAGGAAGAGCTAAAGAGGACAGATACCATATTTCTGGCCAGGTTGTCTC 360
Qy 979 ctggttaaacgctgtggtggaaactgtgctgtgtctcccaattgcaatgaatgtcaa 1038
Db 361 CTGGTCAAGCGCTGTGGAGGAATTTGCCCTGTGTCTCCATATTTGCAATGAATGTCTAG 420
Qy 1039 tgtgtcccaagcaaaagtactataaaataccacagagcttccttcagttgagacccaagacc 1098
Db 421 TGTGTCCACAGTAAAGTTACAAAAAAGTACCATGAGGTCTTTCAGTTGAGACCAAAAACT 480
Qy 1099 ggtgtcagggaggtgacaaatcaactcaccagacgtggccctggagcaccatgagagtg 1158
Db 481 GGAATCAAGGATGTCATAAAGTACTCAATGATGTGCTCTTGAACACACCCAGGAATGT 540
Qy 1159 gactgtgtgtgcagagggagcac 1181
Db 541 GACTGGGTGTAGAGGGGAACGC 563

RESULT 14
BG112388 983 bp mRNA EST 30-JAN-2001
LOCUS 602281957F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4369430 5',
DEFINITION mRNA sequence.
ACCESSION BG112388
VERSION BG112388.1 GI:12605894
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 983)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC <http://mgc.nci.nih.gov/>.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM10025 row: a column: 15
High quality sequence stop: 397.
Location/Qualifiers
1. 983
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4369430"

FEATURES
source

Job time: 4474 sec

7

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29., 2001, 17:05:08 ; Search time 22.8 seconds
(without alignments)
917.337 Million cell updates/sec

Title: US-09-457-066-2
Perfect score: 1858
Sequence: 1 MSFLGILLLTALAGORQGT.....DVALEHHECDVCVRGSGTG 345

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----|-------------|
| 1 | 1858 | 100.0 | 345 | 20 | AA1980 |
| 2 | 1858 | 100.0 | 345 | 20 | AA1981 |
| 3 | 1858 | 100.0 | 345 | 20 | AA1982 |
| 4 | 1858 | 100.0 | 345 | 21 | AA1983 |
| 5 | 1858 | 100.0 | 345 | 21 | AA1984 |
| 6 | 1858 | 100.0 | 345 | 21 | AA1985 |
| 7 | 1858 | 100.0 | 345 | 21 | AA1986 |
| 8 | 1858 | 100.0 | 345 | 21 | AA1987 |
| 9 | 1858 | 100.0 | 345 | 21 | AA1988 |
| 10 | 1858 | 100.0 | 345 | 21 | AA1989 |
| 11 | 1858 | 100.0 | 345 | 21 | AA1990 |

| | | | | | |
|----|--------|-------|-----|----|---------|
| 12 | 1858 | 100.0 | 345 | 21 | AA19578 |
| 13 | 1858 | 100.0 | 345 | 21 | AA19579 |
| 14 | 1858 | 100.0 | 345 | 21 | AA19580 |
| 15 | 1858 | 100.0 | 345 | 21 | AA19581 |
| 16 | 1858 | 100.0 | 345 | 21 | AA19582 |
| 17 | 1858 | 100.0 | 345 | 21 | AA19583 |
| 18 | 1858 | 100.0 | 345 | 21 | AA19584 |
| 19 | 1858 | 100.0 | 345 | 21 | AA19585 |
| 20 | 1858 | 100.0 | 345 | 21 | AA19586 |
| 21 | 1858 | 100.0 | 345 | 21 | AA19587 |
| 22 | 1851 | 99.6 | 345 | 21 | AA19588 |
| 23 | 1850 | 99.6 | 345 | 21 | AA19589 |
| 24 | 1850 | 99.6 | 345 | 21 | AA19590 |
| 25 | 1850 | 99.6 | 345 | 21 | AA19591 |
| 26 | 1745 | 93.9 | 345 | 21 | AA19592 |
| 27 | 1745 | 93.9 | 345 | 21 | AA19593 |
| 28 | 1728 | 93.0 | 318 | 21 | AA19594 |
| 29 | 1671.5 | 90.0 | 339 | 21 | AA19595 |
| 30 | 1667 | 89.7 | 345 | 21 | AA19596 |
| 31 | 1667 | 89.7 | 345 | 21 | AA19597 |
| 32 | 1667 | 89.7 | 345 | 21 | AA19598 |
| 33 | 1450.5 | 78.1 | 282 | 21 | AA19599 |
| 34 | 1316 | 70.8 | 297 | 21 | AA19600 |
| 35 | 814 | 43.8 | 167 | 21 | AA19601 |
| 36 | 798 | 42.9 | 168 | 21 | AA19602 |
| 37 | 753 | 40.5 | 180 | 21 | AA19603 |
| 38 | 742.5 | 40.0 | 370 | 21 | AA19604 |
| 39 | 742.5 | 40.0 | 370 | 21 | AA19605 |
| 40 | 742.5 | 40.0 | 370 | 21 | AA19606 |
| 41 | 742.5 | 40.0 | 370 | 21 | AA19607 |
| 42 | 739.5 | 39.8 | 322 | 21 | AA19608 |
| 43 | 737.5 | 39.7 | 370 | 21 | AA19609 |
| 44 | 735.5 | 39.6 | 370 | 21 | AA19610 |
| 45 | 646 | 34.8 | 149 | 21 | AA19611 |

ALIGNMENTS

RESULT 1
AA19612
ID AA19613 standard; Protein; 345 AA.
AC AA19614;
XX
XX
DT 11-JAN-2000 (first entry)
XX
XX Human VEGF-E protein.
DE
XX VEGF-E; human; vascular endothelial cell growth factor; wound repair;
KW treatment; cardiovascular disorder; endothelial disorder; therapy;
KW tissue generation; regeneration; cardiac hypertrophy; cancer; detection;
KW angiogenic disorder; age-related macular degeneration; vascular disease;
KW neovascularization; tumor; gene mapping.
XX
OS Homo sapiens.
XX
XX WO947677-A2.
XX
XX
PD 23-SEP-1999.
XX
PF 10-MAR-1999; 99WO-0505190.
XX
PR 17-MAR-1998; 98US-0040220.
PR 02-NOV-1998; 98US-0184216.
XX
XX
PA (GETH) GENENTECH INC.
XX
XX Ferrara N, Kuo SS;
PI
XX Human VEGF-X prote
XX Human VEGF-X prote
XX WPI: 1999-580306/49.
DR N-PSDB; AA23691.
XX

PT New growth factor polypeptide useful for treating cardiovascular or
PT endothelial disorders, e.g. cardiac hypertrophy
XX
PS Claim 1; Fig 2; 122pp; English.
XX

CC This invention describes the isolation of a novel human vascular
CC endothelial cell growth factor-E (VEGF-E) polypeptide which has
CC tranquilizer, vulnerary and cardiant activity. VEGF-E can be administered
CC therapeutically, especially by expressing encoding polynucleotides, to
CC treat cardiovascular or endothelial disorders in mammals, especially
CC humans. It is useful in wound repair and tissue generation and
CC regeneration, and may especially be used to treat cardiac hypertrophy
CC It can be combined with a carrier in pharmaceutical compositions, which
CC can be administered to treat disorders as above. VEGF-E can be used to
CC screen for antagonists and agonists, and the antagonists administered to
CC treat angiogenic disorders in mammals (especially humans) e.g. cancer or
CC age-related macular degeneration. It can be used to generate antibodies,
CC useful therapeutically as antagonists, as above. The antibodies are also
CC useful to detect VEGF-E polypeptide, especially to diagnose
CC cardiovascular, endothelial or angiogenic disorders in mammals (e.g.
CC vascular disease, or neovascularization associated with tumor formation),
CC by contacting the antibody with a tissue sample and detecting formation
CC of an antibody-VEGF-E polypeptide complex. Polynucleotides encoding
CC VEGF-E can be used to diagnose cardiovascular and endothelial disorders
CC in mammals, by detecting abnormally high or low VEGF-E gene expression in
CC tissue samples. They can also be used to diagnose a disease or
CC susceptibility to a disease related to a mutated form of VEGF-E (e.g. a
CC cardiovascular, endothelial or angiogenic disorder such as a tumor), by
CC detecting a mutation in the VEGF-E encoding sequence isolated from a
CC sample. They may also be used to produce probes useful to detect related
CC sequences or for gene mapping. This sequence represents the human VEGF-E
CC protein described in the method of the invention.
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 20; Length 345;
Best Local Similarity 100.0%; Pred. NO. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFLGILLTSALAGRGQTAEENLSKQFSSNKQNGVODPQHERILIVSTNGSIHS 60
DB 1 msflgilltsalagrgqtgaenlskqfssnkqngvdpqherilivstngsihs 60
QY 61 PRFHTYPRNTLVRLVAEENWVIOITDERFGLDEPDIDICKYDFVEEPPSDGTIL 120
DB 61 prfhtyprntlvrlvaeenwvioletderfglepdidickdyfveeppsdtgil 120
QY 121 GRWCGSTVPGKQISKGNQIRFVSDEYFPSEPGFCIHYNVMPQFTFVAVSPVLPPSA 180
DB 121 grwcgstvpgkqiskgnqirfvsdeyfpsepgfcihynvmpqftfvaavspvlppsa 180
QY 181 LPDLNNAITATSTLEDLTRYLEPERWQDLEDLYRPTWQLLGAFFVGRKSRVDNLN 240
DB 181 lpdlnnaitatstledltryleperwqddledlyrptwqlqkafvgrksrvvdnl 240
QY 241 LTEEVRLYSTPRNFVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQCVPK 300
DB 241 lteevrlystprnfvsireelkrtdtifwpgcllvkrcgncacclhncnecqcvpk 300
QY 301 VTKKYHEVLQRPKTVGRGLHKSITDVALEHHECDVCVCRSTGG 345
DB 301 vtkeyhevlqrpktrvgrglhksitdvalhehhecdvcvcrstgg 345

RESULT 2

AA41766
ID AA41766 standard; Protein: 345 AA.
XX
AC AA41766;
XX
DT 07-DEC-1999 (first entry)
XX

DE Human PRO200 protein sequence.
XX
KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
OS secreted protein; transmembrane protein.
XX Homo sapiens.
PN WO9946281-A2.
XX
PD 16-SEP-1999.
XX
PF 08-MAR-1999; 99WO-US05028.
XX
PR 10-MAR-1998; 98US-0077450.
PR 11-MAR-1998; 98US-0077632.
PR 11-MAR-1998; 98US-0077641.
PR 11-MAR-1998; 98US-0077649.
PR 12-MAR-1998; 98US-0077791.
PR 13-MAR-1998; 98US-0078004.
PR 17-MAR-1998; 98US-0040220.
PR 20-MAR-1998; 98US-0078886.
PR 20-MAR-1998; 98US-0078910.
PR 20-MAR-1998; 98US-0078936.
PR 20-MAR-1998; 98US-0078939.
PR 25-MAR-1998; 98US-0079294.
PR 26-MAR-1998; 98US-0079656.
PR 27-MAR-1998; 98US-0079663.
PR 27-MAR-1998; 98US-0079664.
PR 27-MAR-1998; 98US-0079689.
PR 27-MAR-1998; 98US-0079728.
PR 27-MAR-1998; 98US-0079786.
PR 30-MAR-1998; 98US-0079920.
PR 30-MAR-1998; 98US-0079923.
PR 31-MAR-1998; 98US-0080105.
PR 31-MAR-1998; 98US-0080107.
PR 31-MAR-1998; 98US-0080165.
PR 31-MAR-1998; 98US-0080194.
PR 01-APR-1998; 98US-0080327.
PR 01-APR-1998; 98US-0080328.
PR 01-APR-1998; 98US-0080333.
PR 01-APR-1998; 98US-0080334.
PR 08-APR-1998; 98US-0081049.
PR 08-APR-1998; 98US-0081070.
PR 09-APR-1998; 98US-0081071.
PR 09-APR-1998; 98US-0081195.
PR 09-APR-1998; 98US-0081203.
PR 09-APR-1998; 98US-0081229.
PR 15-APR-1998; 98US-0081817.
PR 15-APR-1998; 98US-0081838.
PR 15-APR-1998; 98US-0081952.
PR 15-APR-1998; 98US-0081955.
PR 21-APR-1998; 98US-0082568.
PR 21-APR-1998; 98US-0082569.
PR 22-APR-1998; 98US-0082700.
PR 22-APR-1998; 98US-0082704.
PR 23-APR-1998; 98US-0082804.
PR 23-APR-1998; 98US-0082767.
PR 23-APR-1998; 98US-0082796.
PR 27-APR-1998; 98US-0083336.
PR 28-APR-1998; 98US-0083322.
PR 29-APR-1998; 98US-0083392.
PR 29-APR-1998; 98US-0083495.
PR 29-APR-1998; 98US-0083496.
PR 29-APR-1998; 98US-0083499.
PR 29-APR-1998; 98US-0083500.
PR 29-APR-1998; 98US-0083545.
PR 29-APR-1998; 98US-0083554.
PR 29-APR-1998; 98US-0083558.
PR 29-APR-1998; 98US-0083559.
PR 30-APR-1998; 98US-0083742.
PR 05-MAY-1998; 98US-0084366.
PR 06-MAY-1998; 98US-0084414.

```
PR 06-MAY-1998; 98US-0084441.
PR 07-MAY-1998; 98US-0084598.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084627.
PR 07-MAY-1998; 98US-0084637.
PR 07-MAY-1998; 98US-0084639.
PR 07-MAY-1998; 98US-0084640.
PR 07-MAY-1998; 98US-0084643.
PR 13-MAY-1998; 98US-0085323.
PR 13-MAY-1998; 98US-0085338.
PR 13-MAY-1998; 98US-0085339.
PR 15-MAY-1998; 98US-0085573.
PR 15-MAY-1998; 98US-0085579.
PR 15-MAY-1998; 98US-0085580.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085689.
PR 15-MAY-1998; 98US-0085697.
PR 15-MAY-1998; 98US-0085700.
PR 15-MAY-1998; 98US-0085700.
PR 18-MAY-1998; 98US-0086023.
PR 22-MAY-1998; 98US-0086392.
PR 22-MAY-1998; 98US-0086414.
PR 22-MAY-1998; 98US-0086430.
PR 22-MAY-1998; 98US-0086486.
PR 22-MAY-1998; 98US-0087098.
PR 28-MAY-1998; 98US-0087106.
PR 28-MAY-1998; 98US-0087208.
PR 30-JUL-1998; 98US-0094651.
PR 11-SEP-1998; 98US-0100038.
XX
PA (GETH ) GENENTECH INC.
XX
PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX
DR WPI; 1999-551358/46.
DR N-PSDB; AAZ34296.
XX
XX
XX New secreted and transmembrane polypeptides and their polynucleotides,
PT useful for treating blood coagulation disorders, cancers and cellular
PT adhesion disorders -
XX
PS Claim 12; Fig 207; 530pp; English.
XX
XX The present invention describes secreted and transmembrane polypeptides
CC and their polynucleotides. The nucleotide sequences are useful as
CC sources of probes, primers, for chromosome mapping, and for generation
CC of antisense sequences. They can also be used to create transgenic
CC animals. The proteins can be used to treat a variety of diseases and
CC disorders, depending on their function. Diseases that may be treated
CC include blood coagulation disorders, cancers and cellular adhesion
CC disorders. They may also be used to raise antibodies. AAZ33891 to
CC AAZ34338, and AAZ41685 to AAZ41774 represent polynucleotide and
CC polypeptide sequence given in the exemplification of the present
CC invention.
XX
XX Sequence 345 AA;
XX
XX Query Match 100.0%; Score 1858; DB 20; Length 345;
XX Best Local Similarity 100.0%; Pred. No. 6e-180;
XX Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MSLFGLLLTSLAGQROQTAEISLSSKQFSSNKQNGVQDPQHERIITVSTNGSIHS 60
Db 1 mslfigllltsalagrgqtgaesnlsskftssnkegpgvgqdpqheriitvstngsihs 60
QY 61 PRFPHTYPRNTVLVWELVAVENWVLIQLTDFERFGLDEDDICKYDFVEVEPSPDGTIL 120
Db 61 prfphtyprntvlvwrivaveenvwvliqtdferfgledpeddickdyfveveepsdgtil 120
QY 121 GRWCSGTVPGKQISKNGOIRIRFVSDVEYFSPGFCIHYNIVMPQFTAVSPSVLPSSA 180
Db 121 grwcsgtvpgkqiskngqirirfvsdveypspgfcihynivmpqfteavspsvlpssa 180
```

```
QY 181 LPDLLNNAITAFSTLEDLIRYLEPERWOLDLEDLYRPTWOLLGKAFVGRKSRVVDLNL 240
Db 181 lpidllnnaitafstledliryleperwoldledlyrptwollgkafvgrksrvvdlnl 240
QY 241 LTEEVRLYSCTPRNFSVSIREEKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
Db 241 lteevrlyscprnfsvsireekrttdtlfwpgcllvkrcgncaccclhncnecqcvpsk 300
QY 301 VTKKYHEVLQLRPKTVGRGLHKSITDVALEHHEECDCVCRGSTGG 345
Db 301 vtkkyhevlqlrpkrtgvrghksitdvalehheecdcvcrgstgg 345
XX
XX RESULT 3
XX AAY30023
XX ID AAY30023 standard; Protein; 345 AA.
XX AC AAY30023;
XX DT 11-OCT-1999 (first entry)
XX DE Human vascular endothelial growth factor related protein.
XX KW Vascular endothelial growth factor related protein; VEGF-R protein;
XX KW tissue growth inhibition; tumour growth; cancer; tissue growth;
XX KW angiogenesis; coronary artery blockage.
XX OS Homo sapiens.
XX PN WO9937671-A1.
XX PD 29-JUL-1999.
XX PF 26-JAN-1999; 99WO-US01574.
XX PR 31-AUG-1998; 98US-0098548.
XX PR 27-JAN-1998; 98US-0072635.
XX PR 05-JUN-1998; 98US-0088089.
XX PR 24-JUN-1998; 98US-0090544.
XX PA (ELIL ) LILLY & CO ELI.
XX PI Dou S, Na S, Song HY;
XX WPI; 1999-458680/38.
XX DR N-PSDB; AAX86352.
XX PT A vascular endothelial growth factor related protein and related
XX polynucleotide, useful for identifying antagonists and binding
XX compounds
XX Claim 1; Page 56-58; 62pp; English.
XX
XX The present sequence represents a vascular endothelial growth factor
XX related (VEGF-R) protein. VEGF-R can be used in assays to identify
XX compounds that bind to it or that antagonize its activity. VEGF-R
XX antagonists (e.g. anti-VEGF-R antibodies) are useful for inhibiting
XX tissue growth. This is useful for inhibiting tumour growth and for
XX treating cancer. VEGF-R itself can be used to stimulate tissue
XX growth, angiogenesis and to treat coronary artery blockage. The
XX VEGF-R coding sequence can be used for the recombinant production of
XX the VEGF-R protein.
XX
XX Sequence 345 AA;
```

```
Query Match 100.0%; Score 1858; DB 20; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSLFGLLLTSLAGQROQTAEISLSSKQFSSNKQNGVQDPQHERIITVSTNGSIHS 60
XX
```

| | | | |
|----|-----|--|-----|
| Db | 1 | mslfigllltsalagrqrqtqæälnlskfkfsnkegnvgvdpqdheriitvstngsihs | 60 |
| QY | 61 | PREPHTYPRNTVLVWLVAVEENVWLIQTLTDFRFGLEDPEDDICKYDFEVEEPPSDGTIL | 120 |
| Db | 61 | prfphyprrntvlvwlvaaveenvwiqtfderrfgledpeddickdyfveeepsdgtil | 120 |
| QY | 121 | GRWCSGSTVPGKQISKNGNOIRIRFVSDSEYFSPCEGFCITHYNTVMPOQTFEAVSPSVLPESA | 180 |
| Db | 121 | grwcsgstvpvgkqiskngnqirirfvsdseyfspcgfcihynlvmpqfleaavspsvlpesa | 180 |
| QY | 181 | LPDLINNAITAFSTLEDILRYLEPERWQLOLEDILYRPTWOLLGKAFVGRKSRVVDLNL | 240 |
| Db | 181 | lpdlinnaitafstledilryleperwqldedilrptwqlglgkafvgrksrvvdlnl | 240 |
| QY | 241 | LTEEVRLYSCTPRNFVSVIREELKRTDTITFPGCLLVKRCGNCACCLHNCNECQVPSK | 300 |
| Db | 241 | lteevrlyscptrnfsvsireelkrtdtitfpgcllvrcgncacclhncnecqvpsk | 300 |
| QY | 301 | VTKKYHEVILQLRPKTVGRGLHKLSLTDVALEHHEECDCVCRGSTGG | 345 |
| Db | 301 | vtkkyhevliqlrpktvgrglhkslldvalenhheecdcvcrastgg | 345 |

RESULTS

RESOLLI 4
AAB48657
ID AAB48657 standard: Protein: 345 AA.

AA AAB48657:

09-MAR-2001 (first entry)

Human zveaf3, SEO ID NO:33.

Human; zvegf3; zvegf4 fusion; growth factor homologue; VEGF/PDGF family; CUB domain; PDGF-like activity; mitogenic; osteogenic; neovascularisation; tissue repair; proliferation; differentiation; liver damage; neuroregenerative; Alzheimer's disease; multiple sclerosis; periodontal disease; bone fracture; wound healing; vulnery; ischaemia; immunomodulation; hepatic.

OS Homo sapiens.

AA PN WO200066736-A1.

09-NOV-2000

03-MAY-2000: 2000WO-PS40047XX

XX
PR
03-MAY-1999.
99JUS-0304216.

PR IO-NOV-1999; 99US-0164463.
PR 04-FEB-2000; 2000US-0180169.

XX
PA (ZYMO) ZYMOGENETICS INC.

PT Gilbert T. Hart C.F. she

XX
DR
WPT: 2000-687541/67.

DR N-PSDB; AAC81582.
XX

PT Growth factor homologues for treating

| | |
|----|---|
| PT | Alzheimer's disease - |
| XX | |
| XX | |
| PS | Clajm 48: page 125-126: 143nn: English. |

XX The invention relates to the human growth

CC (AA048603) and nucleic acids encoding it (AA048135). Zvegf4 is a member
CC of the PDGF (platelet-derived growth factor)/VEGF (vascular endothelial
CC growth factor) family. Zvegf4 has a growth factor domain (AAB48654)
CC characterised by a PDGF cysteine knot structure, and a CUB domain
CC (AAB48655) which has a beta barrel structure. Zvegf4 has PDGF-like
CC activity, having mitogenic activity on fibroblasts, vascular smooth
CC muscle cells and pericytes, and has also been shown to stimulate bone
CC

growth. The invention also relates to fusion proteins comprising human zvegf4 or fragments thereof, particularly human zvegf4/human zvegf3 fusions; expression constructs and host cells comprising human zvegf4 nucleic acids; the recombinant expression of human zvegf4; an antibody which binds to human zvegf4 or a fragment thereof; a method of activating a cell-surface PDGF receptor using a zvegf4-derived polypeptide; a method of modulating the proliferation, differentiation, migration or metabolism of bone cells, comprising exposing bone cells to zvegf4-derived polypeptides; and a method of detecting a genetic abnormality in the zvegf4 gene of a patient. zvegf4 proteins and derived fragments may be used to stimulate tissue development or repair, or cellular differentiation or proliferation. They are particularly used for the treatment or repair of liver damage, and may also be used to modulate neurite growth (e.g., in the treatment of Alzheimer's disease or multiple sclerosis). Due to their osteogenic activity, they may be used in the treatment of periodontal disease and fractures. They may also be used to enhance expansion and mobilisation of haematopoietic stem cells and endothelial precursor stem cells, which may be useful in the treatment of ischaemia, in wound healing, and in the modulation of the immune system. The present sequence represents human zvegf3.

| | | | | |
|-----------------------|--------------|-------------------|-----------|-------------|
| Query Match | 100.0% | Score 1858; | DB 21; | Length 345; |
| Best Local Similarity | 100.0% | Pred. No. 6e-180; | | |
| Matches 345: | Conservative | 0: Mismatches | 0: Indels | 0: Gaps |

QY 1 MSLEGLLLTSALAGOROGTOAESNLSSKFOFSSNKEONGVODPOHERIITVSTNGSIHS 60

1 mslfqlllltsalagqrqatqaesnlsskffssnkeqngvdpbqheriitvstnqsihs 60

61 PREPHTYPRNTVLVWRLVAVEENVWIOLTFDERRFGLEDPEDDICKYDFEVEEPEPSDGTIL 120

61 prfphtyprntv lvwrlvaveenvwqlt fderfaledpeddickdyfveveepsdqtl 120

QY 121 GRWCGSGTVPGKOISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPOFTEAVSPSVLPSPA 180

Db 121 qrwcsqstvqakaiskqnqirirfvsdevfpsepqfcihynivmpqfteavspsvlppsa 180

QY 181 LPLDLLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLGKAFVFGKRSRVVDLNL 240

Db 181 lpldllnnaitafstledliryleperwqldledlyrptwqllqkafvfqrksrvvdnl 240

QY 241 L TEEVRLYSC TPRNFSVSIREELKRDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSK 300

Db 241 lteevrlyscprnfsvsireelkrttdifwpqcllvkrcqqncacclhncnecqcvpsk 300

QY 301 VTKKYHEVLQLRPKTVGRGLHKS LTDVALEHHEECDVCVRGSGG 345

Db 301 vtkkyhevqlrpkqvrglhksltdvalehheecdvcrgstqq 345

RESULT 5

RESOL
AAB24250
ID AAB24250 standard; Protein; 345 AA.

AAB24250;

DT 08-FEB-2001 (first entry)

DE Human platelet-derived growth factor related protein LP8.

KW Human; platelet derived growth factor related protein; LP8; VEGFh;
KW vascular endothelial growth factor h; tissue regeneration; vulneryary;
KW atherosclerosis; PDGF-related protein; antiarteriosclerotic.

OS Homo sapiens.

PN WO200059940-A2.

12-OCT-2000.
PD

PF 24-MAR-2000; 2000WO-US06427.
XX
PR 06-APR-1999; 99US-0127913.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Hammond LJ, Na S;
XX
DR WPI; 2000-664991/64.
DR N-PSDB; AAC64426.
XX
XX Enhancing tissue growth and promoting wound healing by administering
PT platelet-derived growth factor related protein, LP8 or its analog and
PT treating atherosclerosis by administering LP8 antagonist
XX
XX Claim 4; Page 63-64; 64pp; English.
XX
CC The present invention describes a method for enhancing tissue growth,
CC promoting wound healing or stimulating smooth muscle growth by
CC administering a platelet-derived growth factor (PDGF) related protein,
CC designated LP8 or its analogue. Also described is a method of slowing
CC the progress of atherosclerosis or treating atherosclerosis comprising
CC the administration of an LP8 antagonist. The method is useful for
CC enhancing tissue growth, promoting wound healing and stimulating smooth
CC muscle growth. Antagonists of LP8 are useful for treating
CC atherosclerosis. The present sequence represents human LP8, which is
CC also called VEGFh.
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLFGLLLTSALAGORQGTQAESNLSSKFQFSSNKQNGVQDPQHERIITVSTNGSIHS 60
DB 1 mslfglilltsalagrgtqaesnlsskfssnkqngvqdpqheriitvstngsihs 60

QY 61 PRPHPTYPRNTVLVRLVAEENWVLIQLTDFDERFGLDEPDICKYDFVEEPEPSDGTIL 120
DB 61 prphptyprntvlvrlvaeenwvliqltderfglepddickdyfveveepsdgtil 120

QY 121 GRWCGSGTVPKGQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTAVSPSVLPSPA 180
DB 121 grwcgsgtvpkgqiskgnqirirfvsdeyfpsepgfcihynivmpqfteavspsvlpssa 180

QY 181 LPDLNNAITAFSTLEDLIRYLEPERWOLDLEDLYRPTWOLLGKAFVGRKSRVVDLNL 240
DB 181 lpdlnnaitafstledliryleperwoldledlyrptwollgkafvgrksrvvdlnl 240

QY 241 LTEEVRVLSCTPRNFSVSIREEKRTDITFWPGCLLVKRCGNACCLHNCNECQVPSK 300
DB 241 lteevrlyscprnfsvsireekrtditfwpgcllvkrcgncacclhncnecqvp sk 300

QY 301 VTKKYHEVLQRPKGVGRGLHSLTDVALEHHEBDCVCRGSTGG 345
DB 301 vtkeyhevlqrpktgvrghlskltldvalehheebdcvcrgstgg 345

RESULT 6
AAB44322
ID AAB44322 standard; Protein; 345 AA.
XX
AC AAB44322;
XX
XX
DT 08-FEB-2001 (first entry)
XX
DE Human PRO200 (UNQ174) protein sequence SEQ ID NO:488.
XX
KW Human; secreted protein; transmembrane protein; PRO; EST; cytosolic;
KW expressed sequence tag; detection; cancer.
XX

OS Homo sapiens.
XX
PN WO200053756-A2.
XX
PD 14-SEP-2000.
XX
PF 18-FEB-2000; 2000WO-US04341.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 12-MAR-1999; 99US-0123957.
PR 29-MAR-1999; 99US-0126773.
PR 21-APR-1999; 99US-0130232.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 23-JUN-1999; 99US-0141037.
PR 26-JUL-1999; 99US-0145698.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 30-DEC-1999; 99WO-US31243.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Kljavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;
XX
DR WPI; 2000-611443/58.
DR N-PSDB; AAC78582.
XX
PT Novel PRO polypeptides and polynucleotides used in detection methods,
PT to target bioactive molecules to specific cells, and to modulate
PT cellular activities -
XX
PS Claim 12; Fig 207; 636pp; English.
XX
CC AAC78458 to AAC78599 represent polynucleotide and EST (expressed
CC sequence tag) sequences which encode secreted or transmembrane PRO
CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic
CC activity. The polynucleotides and polypeptides can be used for detecting
CC the presence of PRO polypeptides in samples, for linking bioactive
CC molecules to cells and for modulating biological activities of cells,
CC using the polypeptides for specific targeting. The polypeptide targeting
CC can be used to kill the target cells, e.g. for the treatment of cancers.
CC The polypeptide pairs provide specific targeting of bioactive molecules
CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
CC the isolation of the PRO polynucleotide sequences.
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLFGLLLTSALAGORQGTQAESNLSSKFQFSSNKQNGVQDPQHERIITVSTNGSIHS 60
DB 1 mslfglilltsalagrgtqaesnlsskfssnkqngvqdpqheriitvstngsihs 60

QY 61 PRPHPTYPRNTVLVRLVAEENWVLIQLTDFDERFGLDEPDICKYDFVEEPEPSDGTIL 120
DB 61 prphptyprntvlvrlvaeenwvliqltderfglepddickdyfveveepsdgtil 120

QY 121 GRWCGSGTVPKGQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTAVSPSVLPSPA 180
DB 121 grwcgsgtvpkgqiskgnqirirfvsdeyfpsepgfcihynivmpqfteavspsvlpssa 180

QY 181 LPDLNNAITAFSTLEDLIRYLEPERWOLDLEDLYRPTWOLLGKAFVGRKSRVVDLNL 240
DB 181 lpdlnnaitafstledliryleperwoldledlyrptwollgkafvgrksrvvdlnl 240

QY 241 LTEEVRVLSCTPRNFSVSIREEKRTDITFWPGCLLVKRCGNACCLHNCNECQVPSK 300
DB 241 lteevrlyscprnfsvsireekrtditfwpgcllvkrcgncacclhncnecqvp sk 300

QY 301 VTKKYHEVLQRPKGVGRGLHSLTDVALEHHEBDCVCRGSTGG 345
DB 301 vtkeyhevlqrpktgvrghlskltldvalehheebdcvcrgstgg 345

RESULT 6
AAB44322
ID AAB44322 standard; Protein; 345 AA.
XX
AC AAB44322;
XX
XX
DT 08-FEB-2001 (first entry)
XX
DE Human PRO200 (UNQ174) protein sequence SEQ ID NO:488.
XX
KW Human; secreted protein; transmembrane protein; PRO; EST; cytosolic;
KW expressed sequence tag; detection; cancer.
XX

Db 121 grwcgsvtpgkqkgnqirirfsvdeyfpsepgfcihynivmpqfteavspvlpssa 180
Qy 181 LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLGKAFVGRKSRVVDLNL 240
Db 181 lpldllnnaitafstledliryleperwqldledlyrptwqlgkafvgrksrvvdnl 240
Qy 241 LEEVRLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
Db 241 lteevrlyscprnfsvsireelkrtdtifwpgcllvkrcgncacclhncnecqvpksk 300
Qy 301 VTKKYHEVLQRLPKTGVGRGLHSLTDVALEHHEECDCVCRGSTGG 345
Db 301 vtkeyhevlqrlpkgtgrglhsltdvalehheecdcvcrgstgg 345

RESULT 7

AAB10633
ID AAB10633 standard; Protein; 345 AA.

XX
AC AAB10633;

XX
DT 19-JAN-2001 (first entry)

XX
DE Human RACE generated VEGF-X protein.

XX
KW VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic;
KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
KW venous sore; diabetic ulcer; burns; skin graft growth.

XX
OS Homo sapiens.

XX
PN WO200037641-A2.

XX
PD 29-JUN-2000.

XX
PF 21-DEC-1999; 99WO-US30503.

XX
PR 22-DEC-1998; 98GB-0028377.

XX
PR 18-MAR-1999; 99US-0124967.

XX
PR 08-NOV-1999; 99US-0164131.

XX
PA (JANC) JANSSEN PHARM NV.

XX
PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;

XX
PI Dhanaraj SN, Xu J;

XX
XX WPI: 2000-442669/38.

XX
DR N-PSDB; AAA71951.

XX
PT New vascular endothelial growth factor protein, useful for treating or
PT preventing diseases associated with inappropriate angiogenesis activity
PT such as cancer, rheumatoid arthritis, psoriasis and wounds -

XX
PS Disclosure; Fig 6; 127pp; English.

XX
CC This invention describes a novel vascular endothelial growth factor-X
CC (VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
CC vulnery, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
CC antidiabetic activity and acts as an angiogenesis and vascularization
CC regulator. An antisense molecule of the invention is useful for treating
CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
CC retinopathy by inhibiting angiogenic activity or inappropriate
CC vascularization including formation and proliferation of new blood
CC vessels, growth and development of tissues, tissue regeneration and organ
CC and tissue repair in a subject. The products of the invention are useful
CC for preparing medicaments for treating wounds such as dermal ulcers,
CC pressure sores, venous sores, diabetic ulcers and burns and to promote
CC skin graft growth, tissue repair, proliferation of new blood vessels,
CC tissue regeneration and organ repair by promoting angiogenic activity or
CC vascularization. This sequence represents the RACE generated human VEGF-X

CC protein described in the method of the invention.

XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 21; Length 345;

Best Local Similarity 100.0%; Pred. No. 6e-180; Indels 0; Gaps 0;
Matches 345; Conservative 0; Mismatches 0;

Qy 1 MSLFGLLLTSALAGRQGTQAESNLSSRFQFSSNKQNGVQDPQHERITITVSTNGSIHS 60

Db 1 mslfglLLTSalagrqtgaesnlsskfssnkqngvqdpqheriitvstngsihs 60

Qy 61 PRFPHYTPRNTVLVWRLVAEENVIQLTFDERFGLDEPDDICKYDFVEVEEPSDGTIL 120

Db 61 prfphytpntvlvwrLvaveenvwiqltderfgleddedickdyfveveepsdgtil 120

Qy 121 GRWCGSVTPGKQISKGNQIRIRFVSDVEYFPSEPGFCIHYNIVMPQFTAEVSPVLPSSA 180

Db 121 grwcgsvtpgkqiskgnqirirfvsdeyfpsepgfcihynivmpqfteavspvlpssa 180

Qy 181 LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLGKAFVGRKSRVVDLNL 240

Db 181 lpldllnnaitafstledliryleperwqldledlyrptwqlgkafvgrksrvvdnl 240

Qy 241 LEEVRLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQVPSK 300

Db 241 lteevrlyscprnfsvsireelkrtdtifwpgcllvkrcgncacclhncnecqvpksk 300

Qy 301 VTKKYHEVLQRLPKTGVGRGLHSLTDVALEHHEECDCVCRGSTGG 345

Db 301 vtkeyhevlqrlpkgtgrglhsltdvalehheecdcvcrgstgg 345

RESULT 8

AAB10635
ID AAB10635 standard; Protein; 345 AA.

XX
AC AAB10635;

XX
DT 19-JAN-2001 (first entry)

XX
DE Human VEGF-X protein #1 isolated from clones 4 and 7.

XX
KW VEGF-X; vascular endothelial growth factor; human; vulnery; cytostatic;
KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
KW venous sore; diabetic ulcer; burns; skin graft growth.

XX
OS Homo sapiens.

XX
PN WO200037641-A2.

XX
PD 29-JUN-2000.

XX
PF 21-DEC-1999; 99WO-US30503.

XX
PR 22-DEC-1998; 98GB-0028377.

XX
PR 18-MAR-1999; 99US-0124967.

XX
PR 08-NOV-1999; 99US-0164131.

XX
PA (JANC) JANSSEN PHARM NV.

XX
PI Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;

XX
PI Dhanaraj SN, Xu J;

XX
DR WPI: 2000-442669/38.

XX
DR N-PSDB; AAA71955.

XX
PT New vascular endothelial growth factor protein, useful for treating or
PT preventing diseases associated with inappropriate angiogenesis activity

PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX Disclosure; Fig 9; 127pp; English.
XX
XX This invention describes a novel vascular endothelial growth factor-x
CC (VEGF-X) protein (Ia) and its encoding polynucleotide (Iia) which has
CC vulnary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
CC antidiabetic activity and acts as an angiogenesis and vascularization
CC regulator. An antisense molecule of the invention is useful for treating
CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
CC retinopathy by inhibiting angiogenic activity or inappropriate
CC vascularization including formation and proliferation of new blood
CC vessels, growth and development of tissues, tissue regeneration and organ
CC and tissue repair in a subject. The products of the invention are useful
CC for preparing medicaments for treating wounds such as dermal ulcers, or
CC pressure sores, venous sores, diabetic ulcers and burns and to promote
CC skin graft growth, tissue repair, proliferation of new blood vessels,
CC tissue regeneration and organ repair by promoting angiogenic activity or
CC vascularization. This sequence represents the human VEGF-X protein
CC isolated from clones 4 and 7 described in the method of the invention.
XX
XX Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSLFGLLLTSALAGORGTQAESNLSSKFQSSNKEQNGVQDPQHERITVTSTNGSIHS 60
DB 1 mslfglilltsalagrggtqaesnlsskfssnkeqngvqdpqheriitvtstngsihs 60
QY 61 PRFHTYPRNTVLRVAVENWVQLTFDERFGLDEPDDICKYDFVEEPSDGTIL 120
DB 61 prfhtyprntvrlvaveenvwqltderfgledpeddickdyfveeepsdgtill 120
QY 121 GRWCSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNVMPQFTEAVSPVLPPSA 180
DB 121 grwcsgtvpqkqiskgnqirirfvsdeyfpsepgfcihynvmpqfteavspvlppsa 180
QY 181 LPDLLNNAITAFSTLEDLIRYLEPERWQDLEDLRYPTWQLLGRKAFVGRKSRVVDLNL 240
DB 181 lpdllnnaiftafstledliryleperwqldledlryptwqlqkafvgrksrvvdlnl 240
QY 241 LTEEVLRYSCPTNFSVSIREELKRTDIFWPGLLVKRCGNCACCLHNCNECQVPSK 300
DB 241 lteevrlyscptnfsvsireelkrtidfwpgcllvkrcgncacclhncnecqvpvk 300
QY 301 VTKKYHEVLQRLPKTGVRLGKSLTDVALEHHEEDCVCRGSTGG 345
DB 301 vtckkyhevlqlrpkgtgvrllgksltdvalehheecdvcvrgstgg 345

RESULT 9
AAB10644
ID AAB10644 standard; Protein; 345 AA.
XX
AC AAB10644;
XX
XX 19-JAN-2001 (first entry)
XX Human VEGF-X protein #4.
XX
XX VEGF-X; vascular endothelial growth factor; human; vulnary; cytostatic;
KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
KW venous sore; diabetic ulcer; burns; skin graft growth.
XX
XX Homo sapiens.
XX
XX WO200037641-A2.

XX 29-JUN-2000.
PD
XX
PF 21-DEC-1999; 99WO-US30503.
XX
PR 22-DEC-1998; 98GB-0028377.
PR 18-MAR-1999; 99US-0124967.
PR 08-NOV-1999; 99US-0164131.
XX (JANC) JANSSEN PHARM NV.
PA
XX Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
PI Dhanaraj SN, Xu J;
XX WPI: 2000-442669/38.
DR N-PSDB; AAA71990.
XX
XX New vascular endothelial growth factor protein, useful for treating or
PT preventing diseases associated with inappropriate angiogenesis activity
PT such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX
XX Disclosure: Fig 30B; 127pp; English.
XX
XX This invention describes a novel vascular endothelial growth factor-x
CC (VEGF-X) protein (Ia) and its encoding polynucleotide (Iia) which has
CC vulnary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
CC antidiabetic activity and acts as an angiogenesis and vascularization
CC regulator. An antisense molecule of the invention is useful for treating
CC or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
CC retinopathy by inhibiting angiogenic activity or inappropriate
CC vascularization including formation and proliferation of new blood
CC vessels, growth and development of tissues, tissue regeneration and organ
CC and tissue repair in a subject. The products of the invention are useful
CC for preparing medicaments for treating wounds such as dermal ulcers,
CC pressure sores, venous sores, diabetic ulcers and burns and to promote
CC skin graft growth, tissue repair, proliferation of new blood vessels,
CC tissue regeneration and organ repair by promoting angiogenic activity or
CC vascularization. This sequence represents a human VEGF-X protein
CC described in the method of the invention.
XX
XX Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSLFGLLLTSALAGORGTQAESNLSSKFQSSNKEQNGVQDPQHERITVTSTNGSIHS 60
DB 1 mslfglilltsalagrggtqaesnlsskfssnkeqngvqdpqheriitvtstngsihs 60
QY 61 PRFHTYPRNTVLRVAVENWVQLTFDERFGLDEPDDICKYDFVEEPSDGTIL 120
DB 61 prfhtyprntvrlvaveenvwqltderfgledpeddickdyfveeepsdgtill 120
QY 121 GRWCSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNVMPQFTEAVSPVLPPSA 180
DB 121 grwcsgtvpqkqiskgnqirirfvsdeyfpsepgfcihynvmpqfteavspvlppsa 180
QY 181 LPDLLNNAITAFSTLEDLIRYLEPERWQDLEDLRYPTWQLLGRKAFVGRKSRVVDLNL 240
DB 181 lpdllnnaiftafstledliryleperwqldledlryptwqlqkafvgrksrvvdlnl 240
QY 241 LTEEVLRYSCPTNFSVSIREELKRTDIFWPGLLVKRCGNCACCLHNCNECQVPSK 300
DB 241 lteevrlyscptnfsvsireelkrtidfwpgcllvkrcgncacclhncnecqvpvk 300
QY 301 VTKKYHEVLQRLPKTGVRLGKSLTDVALEHHEEDCVCRGSTGG 345
DB 301 vtckkyhevlqlrpkgtgvrllgksltdvalehheecdvcvrgstgg 345
RESULT 10

AAB10650
ID AAB10650 standard; Protein; 345 AA.
XX
AC AAB10650;
XX
DT 19-JAN-2001 (first entry)
XX
DE Human 990126vegx protein.
XX
KW VEGF-X; vascular endothelial growth factor; human; vulnery; cytotatic;
KW antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
KW angiogenesis regulator; vascularization regulator; cancer; psoriasis;
KW rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
KW tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
KW venous sore; diabetic ulcer; burns; skin graft growth.
XX
OS Homo sapiens.
XX
XX
XX WO200037641-A2.
XX
XX 29-JUN-2000.
XX
XX 21-DEC-1999; 99WO-US30503.
XX
XX 22-DEC-1998; 98GB-0028377.
XX
XX 18-MAR-1999; 99US-0124967.
XX
XX 08-NOV-1999; 99US-0164131.
XX
XX (JANC) JANSSEN PHARM NV.
XX
XX Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
XX PI Dhanaraj SN, Xu J;
XX
XX WPI; 2000-442669/38.
XX
XX
XX New vascular endothelial growth factor protein, useful for treating or
XX preventing diseases associated with inappropriate angiogenesis activity
XX such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX
XX Disclosure; Fig 11; 127pp; English.
XX
XX This invention describes a novel vascular endothelial growth factor-X
XX (VEGF-X) protein (Ia) and its encoding polynucleotide (Iia) which has
XX vulnery, cytotatic, antirheumatic, antiarthritic, antipsoriatic and
XX antidiabetic activity and acts as an angiogenesis and vascularization
XX regulator. An antisense molecule of the invention is useful for treating
XX or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
XX retinopathy by inhibiting angiogenic activity or inappropriate
XX vascularization including formation and proliferation of new blood
XX vessels, growth and development of tissues, tissue regeneration and organ
XX and tissue repair in a subject. The products of the invention are useful
XX for preparing medicaments for treating wounds such as dermal ulcers,
XX pressure sores, venous sores, diabetic ulcers and burns and to promote
XX skin graft growth, tissue repair, proliferation of new blood vessels,
XX tissue regeneration and organ repair by promoting angiogenic activity or
XX vascularization. This sequence represents the human 990126vegx protein
XX used to illustrate the method of the invention.
XX
XX Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180; Indels 0; Gaps 0;
Matches 345; Conservative 0; Mismatches 0;

QY 1 MSFGLLLTSLAGROGTQAEISLRSKFQSSNKEQNGVDPQHERITVTNGSIHS 60
DB 1 MSFGLLLTSLAGROGTQAEISLRSKFQSSNKEQNGVDPQHERITVTNGSIHS 60
QY 61 PRPHTYPTNTVLVLRVAEENNVTLQTFDERFGLDEPDDICKYDFVEVEPSDGTIL 120
DB 61 PRFHTYPTNTVLVLRVAEENNVTLQTFDERFGLDEPDDICKYDFVEVEPSDGTIL 120

QY 121 GRWCGSTVFGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNVMPQFTAEVSPSVLPSPA 180
DB 121 grwcgsgtvpqgkiskgnqirirfvsdeyfpsepgfcihynvmpqftaeavspsvlpssa 180
QY 181 LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLIRPTWOLLGKAFVGRKSRVVDLNL 240
DB 181 lpidlnnaitafstledliryleperwqldledlirptwllgkafvgrksrvvdlnl 240
QY 241 LTEEVRLYSCTPRNFSVSIREEELKRTDTIFWPGCLLVKRCGNCACCLHNCNECOCVPSK 300
DB 241 lteevrlyscptrnfsvsireelkrtdtifwpgcllvkrcgncacclhncnecocvpsk 300
QY 301 VTKYHVLQRLPKTGVGLHKSITDVALEHHBCDCVCRGSTGG 345
DB 301 vtkkyhevlrpkptgvrghksitdvalehhecdvcvrgstgg 345
RESULT 11
AAB10651
ID AAB10651 standard; Protein; 345 AA.
XX
AC AAB10651;
XX
XX 19-JAN-2001 (first entry)
XX
XX Human VEGF-X protein #3.
XX
XX VEGF-X; vascular endothelial growth factor; human; vulnery; cytotatic;
XX antirheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
XX angiogenesis regulator; vascularization regulator; cancer; psoriasis;
XX rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
XX tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
XX venous sore; diabetic ulcer; burns; skin graft growth.
XX
XX Homo sapiens.
XX
XX WO200037641-A2.
XX
XX 29-JUN-2000.
XX
XX 21-DEC-1999; 99WO-US30503.
XX
XX 22-DEC-1998; 98GB-0028377.
XX
XX 18-MAR-1999; 99US-0124967.
XX
XX 08-NOV-1999; 99US-0164131.
XX
XX (JANC) JANSSEN PHARM NV.
XX
XX Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
XX PI Dhanaraj SN, Xu J;
XX
XX WPI; 2000-442669/38.
XX
XX New vascular endothelial growth factor protein, useful for treating or
XX preventing diseases associated with inappropriate angiogenesis activity
XX such as cancer, rheumatoid arthritis, psoriasis and wounds -
XX
XX Claim 72; Fig 12; 127pp; English.
XX
XX This invention describes a novel vascular endothelial growth factor-X
XX (VEGF-X) protein (Ia) and its encoding polynucleotide (Iia) which has
XX vulnery, cytotatic, antirheumatic, antiarthritic, antipsoriatic and
XX antidiabetic activity and acts as an angiogenesis and vascularization
XX regulator. An antisense molecule of the invention is useful for treating
XX or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
XX retinopathy by inhibiting angiogenic activity or inappropriate
XX vascularization including formation and proliferation of new blood
XX vessels, growth and development of tissues, tissue regeneration and organ
XX and tissue repair in a subject. The products of the invention are useful
XX for preparing medicaments for treating wounds such as dermal ulcers,
XX pressure sores, venous sores, diabetic ulcers and burns and to promote
XX skin graft growth, tissue repair, proliferation of new blood vessels,
XX tissue regeneration and organ repair by promoting angiogenic activity or
XX vascularization. This sequence represents the human 990126vegx protein
XX used to illustrate the method of the invention.
XX
XX Sequence 345 AA;

FR. MODIFIED-SILE 201..207

QY 121 GRWCGSGTVPGKQISKGNQIRIRFVSDEYFPSEP GFCIHYNIVMPQFTEAVSPSVLPpSA 18

I

```
QY 181 LPDLNNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLGKAFVGRKSRVVDNL 240
Db 181 lpldlnnaitafstledliryleperwqldledlyrptwqllgkafvgrksrvvdnl 240
QY 241 LTEEVRLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNCCQVPSK 300
Db 241 lteevrlyscprnfsvsireelkrtdtifwpgcllvkrcgncacclhncnccqvpsk 300
QY 301 VTKKYHEVLQRLPKPTGVRGLHKSITDVALEHHECDVCRCSTGG 345
Db 301 vtkeyhevlqlrpkptgvrghlksitdvaalehhecdvcrcstgg 345

RESULT 13
AAB33414
ID AAB33414 standard; Protein; 345 AA.
AC AAB33414;
XX
XX 29-JAN-2001 (first entry)
XX Human PRO200 protein UNQ174 SEQ ID NO:2.
DE
XX Human; immune related disease; diagnosis; antiinflammatory; cardiant;
KW dermatological; antiarthritic; antirheumatic; immunosuppressive;
KW haemostatic; antithyroid; antidiabetic; nootropic; neuroprotective;
KW antianaemic; hepatotropic; virucide; antipsoariatic; anti allergic;
KW antiasthmatic; systemic lupus erythematosus; rheumatoid arthritis;
KW osteoarthritis; spondyloarthopathy; systemic sclerosis; sarcoidosis;
KW idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;
KW systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;
KW autoimmune thrombocytopaenia; immune-mediated renal disease;
KW demyelinating disease; hepatobiliary disease; Whipple's disease;
KW inflammatory bowel disease; gluten-sensitive enteropathy;
KW autoimmune disease; immune-mediated skin disease; allergic disease;
KW immunological disease; transplantation associated disease;
KW graft rejection; graft-versus-host-disease.
XX
OS Homo sapiens.
XX
XX W0200053758-A2.
XX
XX 14-SEP-2000.
XX
XX 02-MAR-2000; 2000WO-US05841.
XX
XX 08-MAR-1999; 99WO-US05028.
XX 10-MAR-1999; 99US-0123618.
XX 12-MAR-1999; 99US-0123957.
XX 23-MAR-1999; 99US-0125775.
XX 12-APR-1999; 99US-0128849.
XX 20-APR-1999; 99WO-US08615.
XX 28-APR-1999; 99US-0131445.
XX 04-MAY-1999; 99US-0132371.
XX 14-MAY-1999; 99US-0134287.
XX 02-JUN-1999; 99WO-US12252.
XX 23-JUN-1999; 99US-0141037.
XX 20-JUL-1999; 99US-0144758.
XX 26-JUL-1999; 99US-0145698.
XX 28-JUL-1999; 99US-0146222.
XX 01-SEP-1999; 99WO-US20111.
XX 08-SEP-1999; 99WO-US20594.
XX 13-SEP-1999; 99WO-US20944.
XX 15-SEP-1999; 99WO-US21090.
XX 15-SEP-1999; 99WO-US21547.
XX 05-OCT-1999; 99WO-US23089.
XX 29-OCT-1999; 99US-0162506.
XX 29-NOV-1999; 99WO-US28214.
XX 30-NOV-1999; 99WO-US28313.
XX 30-NOV-1999; 99WO-US28409.
XX 01-DEC-1999; 99WO-US28301.
XX 01-DEC-1999; 99WO-US28634.
```

```
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30099.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 11-FEB-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
XX
XX (GETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W;
XX Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V;
XX Stewart TA, Tumas D, Watanabe CK, Wood WI, Yan M;
XX
XX WPI; 2000-572271/53.
XX N-PSDB; AAC58579.
XX
XX Sixty four PRO polypeptides, useful in the diagnosis and treatment of
XX immune related disorders, e.g. systemic lupus erythematosus, rheumatoid
XX arthritis, osteoarthritis, thyroiditis and diabetes mellitus -
XX
XX Claim 33; Fig 2; 309pp; English.
XX
XX The present invention describes sixty four human PRO proteins which can
XX be used in the treatment of immune related diseases. The human PRO
XX proteins, anti-PRO antibodies, agonists and antagonists are useful for
XX treating and diagnosing immune related disorders. The disorders are
XX selected from systemic lupus erythematosus, rheumatoid arthritis,
XX osteoarthritis, juvenile chronic arthritis, spondyloarthopathies,
XX systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
XX syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
XX anaemia, autoimmune thrombocytopaenia, thyroiditis, diabetes mellitus,
XX immune-mediated renal disease, demyelinating diseases of the central
XX CC and peripheral nervous systems, hepatobiliary diseases, inflammatory
XX bowel disease, gluten-sensitive enteropathy and Whipple's disease,
XX autoimmune or immune-mediated skin diseases, allergic diseases,
XX immunological diseases of the lung, and transplantation associated
XX diseases including graft rejection and graft-versus-host-disease.
XX AAC58397 to AAC58578 represent PCR primers and hybridisation probes used
XX in the isolation of human PRO sequences. AAC58579 to AAC58642 and
XX AAB33414 to AAB33477 represent human PRO polynucleotide and protein
XX sequences given in the exemplification of the present invention.
XX
XX Sequence 345 AA;
XX
XX Query Match 100.0%; Score 1858; DB 21; Length 345;
XX Best Local Similarity 100.0%; Pred. No. 6e-180;
XX Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSLFGLLLTSALAGROGTOAESNLSSKQFSSNKONGVODPOHERIITVSTNGSIHS 60
Db 1 mslfgllltsalagrgtgtqaesnlsskfqfssnkeqngvqdpqheriitvstngsihs 60
QY 61 PRFPHTYPRNTVLVWRLVAVERNWIQTDFDERFGLEDPEDDICKYDFVEVEEPPSDGTL 120
Db 61 prfphtypntvvlwrlvaveenvwiqtdfderfgledpeddickdyfveveepsdgtll 120
QY 121 GRWCGSGTVPKGQISKGNQIRIRFVSDEYFSPSEPGFCHYNIVMPQFTFVAVSPVLPPSA 180
Db 121 grwcsgstvpkgqiskgnqirirfvsdeyfspsepgfchynivmpqftfaveavspvlppsa 180
QY 181 LPDLNNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWQLLGKAFVGRKSRVVDNL 240
Db 181 lpldlnnaitafstledliryleperwqldledlyrptwqllgkafvgrksrvvdnl 240
QY 241 LTEEVRLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNCCQVPSK 300
XX 241 lteevrlyscprnfsvsireelkrtdtifwpgcllvkrcgncacclhncnccqvpsk 300
```

Db 241 lteevrlyscprnfsaivrelkrttdtlfwpgcllvkrsgnccacclhncnecqcvpsk 300
QY 301 VTKKYHEVLQRLPKTGVLGHLKSLTDVALEHHEECDCVCRGSTGG 345
Db 301 vtckkyhevlqlrpkgtgvrghlksltdvalehheecdcvcrgstg 345

RESULT 14
AAB24412
ID AAB24412 standard; Protein; 345 AA.
XX AAB24412;
XX
XX
XX 07-NOV-2000 (first entry)
XX Human PRO713 protein sequence SEQ ID NO:137.
XX Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;
KW diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;
KW angiogenic; proliferative; cardiant; cardiovascular; antiatherosclerotic;
KW cycostatic; gene therapy; vaccine.
XX Homo sapiens.
XX OS
XX WO200032221-A2.
XX
XX 08-JUN-2000.
XX
XX 30-NOV-1999; 99WO-US28313.
XX 01-DEC-1998; 98WO-US25108.
XX 16-DEC-1998; 98US-0112850.
XX 12-JAN-1999; 99US-0115534.
XX 08-MAR-1999; 99WO-US05028.
XX 12-MAR-1999; 99US-0123957.
XX 28-APR-1999; 99US-0131445.
XX 14-MAY-1999; 99US-0134287.
XX 02-JUN-1999; 99WO-US12252.
XX 23-JUN-1999; 99US-0141037.
XX 20-JUL-1999; 99US-0144758.
XX 26-JUL-1999; 99US-0145698.
XX 01-SEP-1999; 99WO-US20111.
XX 08-SEP-1999; 99WO-US20594.
XX 13-SEP-1999; 99WO-US20944.
XX 15-SEP-1999; 99WO-US21090.
XX 15-SEP-1999; 99WO-US21547.
XX 05-OCT-1999; 99WO-US23089.
XX 29-OCT-1999; 99US-0162506.
XX (GETH) GENENTECH INC.
XX Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ, Goddard A;
PI Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V;
PI Watanabe CK, Williams PM, Wood WI;
XX WPI; 2000-412154/35.
DR N-PSDB; AAA77621.
XX
XX Nucleic acids encoding PRO polypeptides useful for preventing,
PT diagnosing and treating diagnosing a cardiovascular, endothelial or
PT angiogenic disorders in mammals -
XX
XX Claim 72; Fig 50; 315pp; English.
XX
XX The present invention describes nucleic acids encoding PRO polypeptides
CC useful for preventing, diagnosing and treating diagnosing a
CC cardiovascular, endothelial or angiogenic disorder in mammals by
CC modulating cell proliferation, angiogenesis and cardiovascularisation,
CC and for identifying agonists and antagonists of these processes. The
CC nucleic acids and the proteins they encode may be used in the
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate PRO expression such as cardiovascular, endothelial or

CC angiogenic disorders in mammals (e.g. atherosclerosis, cancers and
CC cardiac hypertrophy). For example, the nucleic acids (NCs) and vectors
CC containing them and the PRO polypeptide may be used to treat disorders
CC associated with decreased PRO expression. AAB77510 to AAA77721 and
CC AAB24388 to AAB24435 represent nucleotide and protein sequences used in
CC the exemplification of the present invention.
XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 1858; DB 21; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSIFGLLLTSLAGOROGTOAESNLSSKQFSSNKQNGVQDPOHERIITVSTNGSIHS 60
Db 1 msifglllltsalagrgtqaesnlsskfssnkeqngvqdpqheriitvstngsihs 60
QY 61 PRFPHTYPRNTVLVRLVAEENWVQLTDFDERFGLDEPDDICKYDFVEVEPSPDGTIL 120
Db 61 prfphtyprntvlvrlvaeenwvqltderfgldedpddickdyfvevepsdgtil 120
QY 121 GRWCGSGTVPKGKQISGNQIRIRFVSDEVPFPGFCIHYNIVMQPTEAVSPVLPSPA 180
Db 121 grwcsgstvpkgkqiskgnqirirfvsdeyfpsepgfcihynivmpqfteavspvlpspa 180
QY 181 LPDLINNAITAFSTLEDLIRYLEPERWQDLEDLRYPTWQLLGKAFVGRKSRVVDNL 240
Db 181 lpdlinnaitafstledliryleperwqldledlryptwqllgkafvgrksrvvdnl 240
QY 241 LTEEVRLYSCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQCVPSK 300
Db 241 lteevrlyscprnfsvsireelkrttdtlfwpgcllvkrsgnccacclhncnecqcvpsk 300
QY 301 VTKKYHEVLQRLPKTGVRGLHLSLTDVALEHHEECDCVCRGSTGG 345
Db 301 vtckkyhevlqlrpkgtgvrghlksltdvalehheecdcvcrgstgg 345

RESULT 15

AAB01419
ID AAB01419 standard; Protein; 345 AA.

XX AAB01419;

XX 20-OCT-2000 (first entry)

XX Human TANGO 128.

XX TANGO; 128; 140; 197; 212; 213; 224; 239; modulating agent; asthma;
KW graft versus-host diseases; rheumatoid arthritis; psoriasis;
KW inflammatory bowel disease; septic shock; ulcerative colitis;
KW Crohn's disease; chronic myelogenous leukemia; cancer; liver
KW disease; Hodgkin's disease; osteoarthritis; Lyme's disease;
KW cachexia; autoimmune disease; myasthenia gravis; autoimmune diabetes;
KW systemic lupus erythematosus; transgenic animal; diagnosis;
KW prognosis; prophylactic; therapeutic; human.

XX Homo sapiens.

XX WO200039284-A1.

XX 06-JUL-2000.

XX 23-DEC-1999; 99WO-US31025.

XX 30-DEC-1998; 98US-0223546.

XX (MILL-) MILLENNIUM PHARM INC.

XX Holtzman DA;

XX WPI; 2000-465743/40.

| | |
|----|---|
| DR | N-PSDB; AAA47452. |
| XX | |
| PT | Novel nucleic acid sequences encoding TANGO-128, 140, 197, 212, 213, |
| CC | 224 and 239 polypeptides useful for the treatment of asthma, rheumatoid |
| PT | arthritis, psoriasis and autoimmune diseases |
| PT | |
| PS | Claim 8; Fig 1; 209pp; English. |
| XX | |
| CC | Nucleic acids encoding TANGO polypeptides are useful as modulating |
| CC | agents for regulating cellular processes like asthma, graft |
| CC | versus-host diseases, rheumatoid arthritis, psoriasis, inflammatory |
| CC | bowel disease, septic shock, ulcerative colitis, Crohn's disease, |
| CC | chronic myelogenous leukemia, cancer, liver disease, Hodgkin's |
| CC | disease, osteoarthritis, Lyme's disease, cachexia and autoimmune |
| CC | diseases e.g. myasthenia gravis, autoimmune diabetes and systemic |
| CC | lupus erythematosus. The nucleic acids are also useful for producing |
| CC | transgenic animals and the TANGO polypeptides themselves. Partial |
| CC | TANGO-128, 140, 197, 212, 224, 239 sequences are useful in |
| CC | forensic biology, for diagnostic assays, prognostic assays, |
| CC | pharmacogenomics and for monitoring clinical trials. TANGO |
| CC | polypeptides are suitable for both prophylactic and therapeutic |
| CC | methods for treating a subject at risk of a disorder or having a |
| CC | disorder associated with aberrant TANGO expression. A wide range |
| CC | of cellular disorders can be treated. |
| XX | |
| SQ | Sequence 345 AA; |
| | |
| | Query Match 100.0%; Score 1858; DB 21; Length 345; |
| | Best Local Similarity 100.0%; Pred. No. 6e-180; |
| | Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| | |
| QY | 1 MSLFGILLTTSALAGORQCQTAAESNLSSKKQFSSNKEONGVDDPQHERRITVSTNGSIHS 60 |
| | |
| Dd | 1 mslfgilllttsalagrqrqtgaesnlsskqfssnkeqngvqpqherritvtngsihs 60 |
| | |
| QY | 61 PRPHPTYPRNTVLWRLVAEENVIQLTDFERFLGLEDPEDDICKYDFVEVEEPSDGTIL 120 |
| | |
| Dd | 61 prphptyprntvlwrlvaeenvviqltderflgledpeddickdyfveveepsdgtil 120 |
| | |
| QY | 121 GRWGCGSTVPKGKIQSKGNQIRIFVSDEYPSPGPFCIHNIWMPQTEAVSPSVLPSPA 180 |
| | |
| Dd | 121 grwcgsstvpkgkiskgnqirifvsdeyfpsepgfcihniwmpqfteavspsvlpssa 180 |
| | |
| QY | 181 LPDLNLNATFTASTLEDLRYLEPERSQWLDELRYPTWQLLGKAFVGRKSRVDNL 240 |
| | |
| Dd | 181 lpdllnnatftastleedlryleperwqldiedlyrptwqllgkafrgrksrvvdnl 240 |
| | |
| QY | 241 LITEVRLYSCTPRNFYSVISREELKRTDTIEWPGCLLYKRCGGNCACCLHNCNECQCVPFSK 300 |
| | |
| Dd | 241 lteevrlysctprnfysvisreelkrtdtiwpgccllykrcggncacclhncneqcqvpsk 300 |
| | |
| QY | 301 VTKKYHEVLRLPKTVGRGLHKSLTDVALHHBECDCVCRGSTGG 345 |
| | |
| Dd | 301 vtkkyhevlrlpktvgvrglhksltdvalehhbecdcvcrgstgg 345 |

Search completed: June 29, 2001, 17:07:42
Job time: 154 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2001, 17:06:18 ; Search time 16.85 seconds
(without alignments)
412.457 Million cell updates/sec

Title: US-09-457-066-2

Perfect score: 1858

Sequence: 1 MSFLGLLLTSLAQRQGT.....DVALEHHHECDVCVRGTGG 345

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|------------------|
| 1 | 169 | 9.1 | 788 | 1 | US-08-572-225-1 |
| 2 | 163.5 | 8.8 | 909 | 3 | US-08-936-135-18 |
| 3 | 163.5 | 8.8 | 926 | 3 | US-08-936-135-20 |
| 4 | 160.5 | 8.6 | 923 | 3 | US-08-936-135-6 |
| 5 | 159.5 | 8.6 | 901 | 3 | US-08-936-135-22 |
| 6 | 159.5 | 8.6 | 906 | 3 | US-08-936-135-24 |
| 7 | 159.5 | 8.6 | 909 | 3 | US-08-936-135-8 |
| 8 | 159.5 | 8.6 | 909 | 3 | US-08-936-135-10 |
| 9 | 159.5 | 8.6 | 914 | 3 | US-08-936-135-12 |
| 10 | 159.5 | 8.6 | 926 | 3 | US-08-936-135-14 |
| 11 | 159.5 | 8.6 | 931 | 3 | US-08-936-135-16 |
| 12 | 154 | 8.3 | 449 | 2 | US-08-839-008-2 |
| 13 | 154 | 8.3 | 449 | 2 | US-08-839-008-9 |
| 14 | 153 | 8.2 | 415 | 4 | US-09-032-523-2 |
| 15 | 149 | 8.0 | 591 | 3 | US-08-991-408-4 |
| 16 | 149 | 8.0 | 1013 | 2 | US-08-866-650-5 |
| 17 | 149 | 8.0 | 1013 | 2 | US-09-021-287-5 |
| 18 | 149 | 8.0 | 1013 | 3 | US-08-991-408-2 |
| 19 | 148 | 8.0 | 1013 | 2 | US-08-866-650-3 |
| 20 | 148 | 8.0 | 1013 | 2 | US-09-021-287-3 |
| 21 | 147 | 7.9 | 354 | 4 | US-08-915-795-5 |
| 22 | 142 | 7.6 | 325 | 4 | US-08-915-795-3 |
| 23 | 140 | 7.5 | 321 | 4 | US-08-915-795-8 |
| 24 | 140 | 7.5 | 358 | 4 | US-08-915-795-8 |
| 25 | 132 | 7.1 | 419 | 2 | US-08-999-811-2 |
| 26 | 132 | 7.1 | 419 | 3 | US-09-042-105-2 |
| 27 | 132 | 7.1 | 419 | 3 | US-09-042-105-18 |

Sequence 8, Appli
Sequence 35, Appli
Sequence 5, Appli
Sequence 7, Appli
Sequence 8, Appli
Sequence 2, Appli
Sequence 11, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 6, Appli
Sequence 6, Appli
Patent No. 5175255
Sequence 4, Appli
Sequence 2, Appli
Sequence 4, Appli
Sequence 33, Appli

US-08-795-430-8
US-08-510-133A-35
US-08-839-008-5
US-08-839-008-7
US-09-032-523-8
PCT-US96-09001-2
US-08-795-430-11
US-08-024-868-2
US-08-242-097-2
US-09-206-695-2
PCT-US96-11995-1
US-08-242-097-6
US-09-206-695-6
5175255-9
US-08-999-811-4
US-08-824-996-2
US-09-042-105-4
US-08-510-133A-33

ALIGNMENTS

RESULT 1
US-08-572-225-1
; Sequence 1, Application US/08572225
; Patent No. 5807981
; GENERAL INFORMATION:
; APPLICANT: Prockop, Darwin J.
; APPLICANT: Hojima, Yoshio
; APPLICANT: Li, Shi-Wu
; APPLICANT: Sieron, Aleksander
; APPLICANT: Brenner, Mitch
; TITLE OF INVENTION: RECOMBINANT C-PROTEINASE AND ITS USE FOR
; TITLE OF INVENTION: DRUG DEVELOPMENT FOR THE TREATMENT OF DISEASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennite & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 13-DEC-1995
APPLICATION NUMBER: US/08/572,225
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: 8389-031
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-3660
TELEFAX: 415-854-3694
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 788 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-572-225-1

Query Match 9.1%; Score 169; DB 1; Length 788;
Best Local Similarity 39.4%; Pred. No. 1.4e-09;
Matches 43; Conservative 17; Mismatches 39; Indels 10; Gaps 5;

RESULT 3
US-08-936-135-20
: Sequence 20, Application US/08936135

```

RESULT      4
US-08-936-135-6
; Sequence 6, Application US/08936135
; Patent No. 6054293
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; APPLICANT: Chen, Hang
; TITLE OF INVENTION: Semaphorin F
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOG
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA

```


ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.30, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,135
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 923 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-6

Query Match 8.6%; Score 160.5; DB 3; Length 923;
Best Local Similarity 36.5%; Pred. No. 1.5e-08;
Matches 42; Conservative 16; Mismatches 50; Indels 7; Gaps 4;
QY 50 INVSTNGSIHSFRPHYTPRNTVLVRLVAEENVVQLTFDERFGLDEPDDICKYDFV 109
Db 31 IKIENPYTSCIPHSYHPSEKCEWLIQAPEYQRIINFNPHFDEKRD---CRYDYV 87
QY 110 EV--EPPSDGTILGRWCGSGTVPGKQISKGNQIRIFVSDYPPSPGFCIHNYI 162
Db 88 EVIDGENEGRLWGKFCGR-IAPSPVSSGPFIFKIFVSD-YETHGAGFSIRYEI 140

RESULT 5
US-08-936-135-22
Sequence 22, Application US/08936135
Patent No. 6054293
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: He, Zhigang
APPLICANT: Chen, Hang
TITLE OF INVENTION: Semaphorin Receptors
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,135
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 901 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-22

Query Match 8.6%; Score 159.5; DB 3; Length 901;
Best Local Similarity 33.1%; Pred. No. 1.8e-08;
Matches 44; Conservative 17; Mismatches 65; Indels 7; Gaps 4;
QY 32 FSSNKEQNGVQDPQHERLIITVSTNGSIHSFRPHYTPRNTVLVRLVAEENVVQLTFD 91
Db 14 YFSGHEVRSQDDPPCGGRPNKSDAGYITSPGYPDYPSHQNCWEIVYAPEPNQKIVLNFN 73
QY 92 ERFGLDEPDDICKYDFVEVE--EPPSDGTILGRWCGSGTVPGKQISKGNQIRIFVSDY 149
Db 74 PHEIEKHD---CKYDFIEIRDGSESADLLGKHCNG-IAPPTIISSGSLVYIKFTSD-Y 128
QY 150 FPSEPGFCIHNYI 162
Db 129 ARQGAGFSIRYEI 141

RESULT 6
US-08-936-135-24
Sequence 24, Application US/08936135
Patent No. 6054293
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: He, Zhigang
APPLICANT: Chen, Hang
TITLE OF INVENTION: Semaphorin Receptors
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,135
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 906 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-24

Query Match 8.6%; Score 159.5; DB 3; Length 906;
Best Local Similarity 33.1%; Pred. No. 1.9e-08;
Matches 44; Conservative 17; Mismatches 65; Indels 7; Gaps 4;

TOPLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-16

Query Match 8.6%; Score 159.5; DB 3; Length 931;
Best Local Similarity 33.1%; Pred. No. 1.9e-08;
Matches 44; Conservative 17; Mismatches 65; Indels 7; Gaps 4;
QY 32 FSNKEQGVQDQHERIITVSTNGSIHPRFPHTYPRNTVLVRLVAEENWVLIQTFD 91
Db 14 YFSGEVRSQDDPCGGRPNKDAGYTSFCYPQDYSHQNCWEIVVAPENQKIVLNF 73
QY 92 ERGLEDPEDDICKYDFVEVE--EPSGDTILGRWCGSGTVPGKQISGNQIRIRFVSDEY 149
Db 74 PHEIEKHD--CKYDFIEIRDGDSADLLGKHGCGN-IAPPTIISGSLVIKFTSD-Y 128
QY 150 FPSEPGFCIHVNI 162
Db 129 AROGAGFSLRYEI 141

RESULT 12
US-08-839-008-2
; Sequence 2, Application US/08839008
; Patent No. 5916758
; GENERAL INFORMATION:
; APPLICANT: Hurlle, Mark R
; APPLICANT: McDonnell, Peter C
; APPLICANT: McNulty, Dean E
; APPLICANT: Rosen, Craig A
; APPLICANT: Siemens, Ivo R
; APPLICANT: Young, Peter R
; APPLICANT: Yue, Tian-Li
; TITLE OF INVENTION: Smooth Muscle Cell-Derived Migration Factor
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/839,008
FILING DATE: 23-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: 08/563,697
FILING DATE: 28-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: P50384
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-839-008-2

Query Match 8.3%; Score 154; DB 2; Length 449;

Best Local Similarity 33.8%; Pred. No. 2.6e-08;
Matches 48; Conservative 22; Mismatches 50; Indels 22; Gaps 10;
QY 56 GSIHSPRPHT-YPRNTVLVRLVAEENWVLIQTFDERFGLDEDDICKYDFVEV--- 111
Db 168 GTLTTPNPESDYPPGISCSWHIIAPPDQV-IALTFF-EKFDLE--PDTYCRDVSVSFNG 223
QY 112 EEPDGGTILGRWCGSGTVPGKQISGNQIRIRFVSDEYPPSPGFCIHVNIWMPQT--- 168
Db 224 AVSDSRRLGKFCGD-AVPGSISSEGNELLVQFVSDLSVTAD-GFSASYK-TLPRGTAK 280
QY 169 -----EAVSPSV-LPDSALP 182
Db 281 GCGPGPKRGTEPKVKLPKRSQP 302

RESULT 13
US-08-839-008-9
; Sequence 9, Application US/08839008
; Patent No. 5916758
; GENERAL INFORMATION:
; APPLICANT: Hurlle, Mark R
; APPLICANT: McDonnell, Peter C
; APPLICANT: McNulty, Dean E
; APPLICANT: Rosen, Craig A
; APPLICANT: Siemens, Ivo R
; APPLICANT: Young, Peter R
; APPLICANT: Yue, Tian-Li
; TITLE OF INVENTION: Smooth Muscle Cell-Derived Migration Factor
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/839,008
FILING DATE: 23-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: 08/563,697
FILING DATE: 28-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: P50384
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-839-008-9

Query Match 8.3%; Score 154; DB 2; Length 449;
Best Local Similarity 33.8%; Pred. No. 2.6e-08;
Matches 48; Conservative 22; Mismatches 50; Indels 22; Gaps 10;

QY 56 GSIHSPRPHT-YPRNTVLVRLVAEENWVLIQTFDERFGLDEDDICKYDFVEV--- 111
Db 168 GTLTTPNPESDYPPGISCSWHIIAPPDQV-IALTFF-EKFDLE--PDTYCRDVSVSFNG 223

QY 112 EPPSOGTILGRWCGSGTVPKGQISGNQIRIRFVSDYFPPSGFCIHYNIVMPOFT--- 168
; Sequence 2, Application US/09032523
; Patent No. 6232454
Db 224 AVSDSRRLGRKFGD-AVPGSISSEGNELLVQFVSDLSVTAD-GFSASYK-TLPRCTAKE 280
QY 169 -----EAVSPSV-LPPSALP 182
; Sequence 4, Application US/08991408
; Patent No. 6008017
Db 281 GQGPGRKRGTEPKVKLPPKSQP 302
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl
; APPLICANT: Baugh, Mariah
; TITLE OF INVENTION: HUMAN PROTEINASE MOLECULES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,523
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0479 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: RATNOT02
; CLONE: 947429
; US-09-032-523-2
Query Match 8.2%; Score 153; DB 4; Length 415;
Best Local Similarity 30.8%; Pred. No. 2.9e-08;
Matches 44; Conservative 25; Mismatches 50; Indels 24; Gaps 8;
QY 55 NGSISHPREP-HTYPRNTVLVRLVAVENWVQLTDFERFGLDEPDICKYDFVEV-- 111
Db 162 SGSFPTNPWRDYPAGVTCVWHIVA-PKNQILKLF-EKFDVE--RDNYCRYDYVAVFN 217
QY 112 -EPPSDGTILGRWCGSGTVPKGQISGNQIRIRFVSDYFPPSGFCIHYNIVMPOFTEA 170
Db 218 GGEVNDARRIGKYCGD-SPPAPIVSENNELLQFLSLTAD-GFIGHY----- 265

QY 171 VSPSVLPPSALPLDLLNNAITAF 193
; Sequence 4, Application US/08991408
; Patent No. 6008017
Db 266 ----IFRPKKLPTTTEQPVTTF 284
; GENERAL INFORMATION:
; APPLICANT: ARLETH, ANTHONY J.
; APPLICANT: WILLETTTE, ROBERT N.
; APPLICANT: ELISHOURBAGY, NABIL A.
; APPLICANT: LI, XIAOTONG
; TITLE OF INVENTION: HUMAN CARDIAC/BRAIN TOLLOID-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,408
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/034,471
; FILING DATE: 02-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: ATG-50038
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-991-408-4
Query Match 8.0%; Score 149; DB 3; Length 591;
Best Local Similarity 34.9%; Pred. No. 1.4e-07;
Matches 38; Conservative 23; Mismatches 38; Indels 10; Gaps 5;
QY 55 NGSISHPRPHTYPRNTVLVRLVA-VEENWVQLTDFERFGLDEPDICKYDFVEVEE 113
Db 204 NGTITPGWPKYPPNKNVCVQVAPTOYRISVKFEFELEG-----NEVCYDYVEIWS 258
QY 114 --PSDGTILGRWCGSGTVPKGQISGNQIRIRFVSDYFPPSGFCIH 160
Db 259 GLSESSEKLHGKCGA-EVPEVITSQFNNMRTFKSDNTV-SKKGKKAHF 305
Search completed: June 29, 2001, 17:09:16
Job time: 178 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2001, 17:06:48 ; Search time 21.63 Seconds
(without alignments)
1214.989 Million cell updates/sec

Title: US-09-457-066-2

Perfect score: 1858

Sequence: 1 MSLLFGLLLTSALAGORQT.....DVALEHHEDCVCRGSTGG 345

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 183.5 | 9.9 | 3623 | 2 T09456 | intrinsic factor-B |
| 2 | 181.5 | 9.8 | 730 | 1 BKH01 | procollagen C-endo |
| 3 | 179.5 | 9.7 | 927 | 1 J00948 | A5 antigen precurs |
| 4 | 176 | 9.5 | 707 | 2 J02218 | procollagen C-endo |
| 5 | 173 | 9.3 | 823 | 1 A58788 | procollagen C-endo |
| 6 | 171 | 9.2 | 3623 | 2 T08618 | intrinsic factor-B |
| 7 | 169 | 9.1 | 986 | 1 B58788 | procollagen C-endo |
| 8 | 169 | 9.1 | 991 | 2 I49540 | procollagen C-endo |
| 9 | 154 | 8.3 | 449 | 2 A55362 | procollagen I C-pr |
| 10 | 148.5 | 8.0 | 1057 | 1 A39288 | dorsal-ventral pat |
| 11 | 143.5 | 7.7 | 686 | 1 A59271 | Ra-reactive factor |
| 12 | 139.5 | 7.5 | 1070 | 2 T31069 | tolluid-BMP-I like |
| 13 | 138.5 | 7.5 | 597 | 2 S71352 | metalloproteinase |
| 14 | 138 | 7.4 | 705 | 1 C1HURB | complement subcomp |
| 15 | 137.5 | 7.4 | 699 | 1 I54763 | Ra-reactive factor |
| 16 | 137.5 | 7.4 | 1524 | 2 T30337 | polyprotein - Afri |
| 17 | 133 | 7.2 | 1594 | 2 T30549 | hensin - rabbit |
| 18 | 132 | 7.1 | 419 | 2 S69207 | vascular endotheli |
| 19 | 130.5 | 7.0 | 1464 | 2 S58984 | development protei |
| 20 | 128 | 6.9 | 402 | 2 JH0403 | procollagen I C-pr |
| 21 | 127.5 | 6.9 | 767 | 2 T30018 | hypothetical prote |
| 22 | 127.5 | 6.9 | 3871 | 2 T22812 | hypothetical prote |
| 23 | 125 | 6.7 | 198 | 2 J50735 | platelet-derived g |
| 24 | 123.5 | 6.6 | 277 | 2 A41735 | hyaluronate-bindin |
| 25 | 120.5 | 6.5 | 245 | 1 TVCTSS | platelet-derived g |
| 26 | 120.5 | 6.5 | 275 | 2 JC6506 | tumor necrosis fac |
| 27 | 118.5 | 6.4 | 276 | 2 A47290 | TSG-6 homolog p34 |
| 28 | 114.5 | 6.2 | 148 | 2 D49530 | 16K vascular endot |
| 29 | 114.5 | 6.2 | 241 | 1 PFHUG2 | platelet-derived g |

| | | | | | |
|----|-------|-----|------|----------|--------------------|
| 30 | 112.5 | 6.1 | 200 | 2 I51551 | platelet-derived g |
| 31 | 112.5 | 6.1 | 215 | 2 S08220 | platelet-derived g |
| 32 | 112.5 | 6.1 | 226 | 2 I51550 | platelet-derived g |
| 33 | 111.5 | 6.0 | 319 | 2 I51569 | UWS-2 protein - Af |
| 34 | 110 | 5.9 | 166 | 2 JN0248 | platelet-derived g |
| 35 | 109 | 5.9 | 197 | 2 S25096 | platelet-derived g |
| 36 | 108.5 | 5.8 | 226 | 1 TMVSS | PDGF-related trans |
| 37 | 108 | 5.8 | 196 | 2 A37359 | platelet-derived g |
| 38 | 104.5 | 5.6 | 232 | 2 A41551 | vascular endotheli |
| 39 | 104 | 5.6 | 196 | 2 B28964 | platelet-derived g |
| 40 | 104 | 5.6 | 211 | 1 PFHUG1 | platelet-derived g |
| 41 | 102.5 | 5.5 | 133 | 2 B49530 | vascular endotheli |
| 42 | 102.5 | 5.5 | 190 | 2 S52130 | vascular endotheli |
| 43 | 102.5 | 5.5 | 196 | 2 A48851 | platelet-derived g |
| 44 | 102.5 | 5.5 | 1023 | 2 G96509 | protein F27F5.21 l |
| 45 | 102.5 | 5.5 | 2083 | 2 T42721 | CRP-ductin-alpha p |

ALIGNMENTS

RESULT 1

T09456

intrinsic factor-B12 receptor Cubilin precursor - human

C:Species: Homo sapiens (man)

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: T09456

R:Kozuyaki, R.; Kristiansen, M.; Silahtaroglu, A.; Hansen, C.; Jacobsen, C.; Tommerup

Blood 91, 3593-3600, 1998

A:Title: The human intrinsic factor-vitamin B12 receptor, cubilin: Molecular character

ion.

A:Reference number: Z16677; MUID:98241400

A:Accession: T09456

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-3623 <KO2>

A:Cross-references: EMBL:AF034611; NID:g3929528; PIDN:AA82612.1; PID:g3929529

C:Genetics:

A:Map position: 10p12

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: receptor; vitamin B12 uptake

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-3623/Product: intrinsic factor-B12 receptor #status predicted <MAT>

F:436-467/Domain: EGF homology <EGF>

Query Match 9.9%; Score 183.5; DB 2; Length 3623;
Best Local Similarity 25.7%; Pred. No. 6.2e-07;
Matches 98; Conservative 46; Mismatches 142; Indels 95; Gaps 24;

QY 10 TSALAGORQ-----GTOAESNLSSKTFQ-----SSNKEQNG-----VODPOHERIIT 51

Db 878 SSSILGSPENKKYCGTDIPSFITSVNFVTVFKSSSTENHGFMKFAEDLACGEILT 937

QY 52 VSTNGSIHSRPFHTYPRNTVLVRLVAEENWVLIQTFDERFCELEDPEDDICKYDFEV 111

Db 938 EST-GTIQSPGHVNVPHGINCTWHIL-VQPNHLIHLMF-ETFHLEPHYN-CTNDYLEV 992

QY 112 EPSPDGTILGRWCGSGTVPGKQISKNQIRIRVSDVEYFPSEPGFCIHYNIV-----MP 165

Db 993 YDTSSETSLGRYCGK-SIPPSLTSSGNSLMLVFVTVSDLAYE-GFLINYAISATAFLQ 1050

QY 166 QTEAVSPVLPSPALPDLNLNNAITAFSTLEDLIRYLEPERMOLDLEDLYR-----PTWQL 222

Db 1051 DVTDDLGTFTSP-----NFPNN-----YNNW-----ECIVRIVRTGQL 1085

QY 223 LGKAFVGRKSRVVDNLITE--EVR-----LY--SCTPRNFVSIREELK-R 265

Db 1086 IAVHFTNFSLEEAIG-NYTFDFLEIRGGYKESPLIGFYGSNLPPTIISHSNKLWLKFK 1144

QY 266 TDTI-----FWPCLLVKRGGNCCACLNHCNEQCQVPSKVTKYHE---VLQLRPK 314

Db 1145 SDQIDTRSGFSAYWDGS--STGCGGN----LTTSSGTFISPNYPMYPYHSSECVWLKSS 1198

QY 315 TGVRLGLKSLTDVLEHHHEEC 335
Db 1199 HG-SAFELEFKDFLHHPNC 1218

RESULT 2
BMH01
procollagen C-endopeptidase (EC 3.4.24.19) precursor, splice form BMP1 - human
N/Alternate names: bone morphogenic protein 1 (BMP1)
C/Species: Homo sapiens (man)
C/Date: 16-Sep-1992 #sequence_revision 03-Aug-1995 #text_change 18-Jun-1999
C/Accession: A37278; E58788
R/Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitsock, L.M.; Whitters, M.J.; Kriz, R.W.; Hew
Science 242, 1528-1534, 1988
A/Title: Novel regulators of bone formation: molecular clones and activities.
A/Reference number: A37278; MUID:89072730
A/Accession: A37278
A/Molecule type: mRNA
A/Residues: 1-730 <WOZ>
A/Cross-references: GB:M22488; NID:gl79499; PIDN:AAA51833.1; PID:gl79500
C/Genetics:
A/Genes: GDB:BMP1
A/Cross-references: GDB:125203; OMIM:112264
A/Map position: 8p21-8p21
C/Function:
A/Description: catalyzes hydrolysis of the carboxyl-terminal propeptide of collagen type
C/Superfamily: procollagen C-endopeptidase; astacin homology; C1r/C1s repeat homology; F
C/Keywords: alternative splicing; beta-hydroxyasparagine; bone; calcium; duplication; gl
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-730/Product: procollagen C-endopeptidase splice form BMP1 #status predicted <MAT>
F:130-321/Domain: astacin homology <AST>
F:322-431/Domain: C1r/C1s repeat homology <C1R1>
F:435-544/Domain: C1r/C1s repeat homology <C1R2>
F:551-587/Domain: EGF homology <EGF>
F:591-700/Domain: C1r/C1s repeat homology <C1R3>
F:91,142,332,363,599/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:163-319,185-205,322-348,375-397,435-461,488-510,551-563,599-572,574-587,591-617,644-66
F:213,217,223,272/Binding site: zinc (His, His, His, Tyr) #status predicted
F:214/Active site: Glu #status predicted
F:565/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 9.8%; Score 181.5; DB 1; Length 730;
Best Local Similarity 37.3%; Pred. NO. 1.2e-07;
Matches 50; Conservative 20; Mismatches 47; Indels 17; Gaps 7;

QY 55 NGSISHPRPHTYPRNTVLVRLVA-VEENVWQLTFDERFGLDEPDDICKYDFVEVE 113
Db 599 NGSITSPGWPKEYPPNKNKINQVAPQYRISIQDFEFTEG-----NDVCKYDFVEVRS 653

QY 114 --PSDGTILGRWCGSGTVPGKQISKGNQIRIRFVSDYFPPSEPGFCIHYNLVMPQFTEAV 171
Db 654 GLTADSKLHGKFCGS-EKPEVITSQIAHMRVEFKSDNTV-SKGGFRAHF-----FSEK- 704

QY 172 SPSVLPSPALPLDL 185
Db 705 RPALQPPRGRPHQL 718

RESULT 3
JQ0948
A5 antigen precursor - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C/Accession: JH0466; JQ0948
R/Takagi, S.; Hirata, T.; Agata, K.; Mochii, M.; Eguchi, G.; Fujisawa, H.
Neuron 7, 295-307, 1991
A/Title: The A5 antigen, a candidate for the neuronal recognition molecule, has homology
A/Reference number: JH0466; MUID:91337458
A/Accession: JH0466
A/Molecule type: mRNA
A/Residues: 1-927 <TAK>

A/Cross-references: GB:D10467; GB:D01077; NID:g222962; PIDN:BA01260.1; PID:g222963
A/Experimental source: tadpole, brain
A/Note: This protein has motifs homologous to complement components C1r and C1s and t
C/Comment: This protein is a neuronal cell surface molecule involved in the neuronal
C/Superfamily: Xenopus A5 antigen; C1r/C1s repeat homology; discoidin I amino-termina
C/Keywords: duplication; glycoprotein; transmembrane protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-927/Product: A5 antigen #status predicted <ASA>
F:27-138/Domain: C1r/C1s repeat homology <C1R1>
F:147-262/Domain: C1r/C1s repeat homology <C1R2>
F:274-424/Domain: discoidin I amino-terminal homology <DN1>
F:430-584/Domain: discoidin I amino-terminal homology <DN2>
F:646-812/Domain: MAM homology <MAM>
F:861-883/Domain: transmembrane #status predicted <TMM>
F:150,261,300,523,844/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.7%; Score 179.5; DB 1; Length 927;
Best Local Similarity 31.6%; Pred. NO. 2.4e-07;
Matches 55; Conservative 29; Mismatches 73; Indels 17; Gaps 7;

QY 50 ITVSTNGSIHSPRPHTYPRNTVLVRLVAVEENVWQLTFDERFGLDEPDDICKYDFV 109
Db 31 IKITSPSVLTSAGYPHSYPPSORCEWLQAPHYQIRIMFNPNPHEDLEDE---CKYDIV 87

QY 110 EV--EEPSDGTILGRWCGSGTVPGKQISKGNQIRIRFVSDYFPPSEPGFCIHYNVMP-- 165
Db 88 EVIDGDNANGQLLKCYCGK-IAPSPVSTGPSIFIRFVSDYETPG-AGFSIRYEVFKTGP 145

QY 166 ---QFTEA--VSPSLPPSALPLDLNNAITAFSTLEDLIRLPERWOLDLE 213
Db 146 ECSRNTSSNGVIKSPKPYKPNALCTYIFAPKMQEIV--LEFSEFELEAD 197

RESULT 4
JC2218
procollagen C-endopeptidase (EC 3.4.24.19) - African clawed frog (clone 22AN)
N/Alternate names: bone morphogenic protein 1
C/Species: Xenopus laevis (African clawed frog)
C/Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Sep-1999
C/Accession: JC2218
R/Maeno, M.; Xue, Y.; Wood, T.I.; Ong, R.C.; Kung, H.
Gene 134, 257-261, 1993
A/Title: Cloning and expression of cDNA encoding Xenopus laevis bone morphogenetic pr
A/Reference number: JC2218; MUID:94085787
A/Accession: JC2218
A/Molecule type: mRNA
A/Residues: 1-707 <MAE>
A/Cross-references: GB:L12249; NID:g406540; PIDN:AAA16313.1; PID:g406541
C/Comment: This protein induces ectopic cartilage formation in vivo.
C/Superfamily: procollagen C-endopeptidase; astacin homology; C1r/C1s repeat homology
C/Keywords: beta-hydroxyasparagine; glycoprotein; hydrolase; metalloproteinase; zinc
F:93-284/Domain: astacin homology <AST>
F:285-397/Region: complement lr/1s-like repeat
F:285-394/Domain: C1r/C1s repeat homology <C1R1>
F:398-510/Region: complement lr/1s-like repeat
F:398-507/Domain: C1r/C1s repeat homology <C1R2>
F:514-550/Domain: EGF homology <EGF>
F:554-666/Region: complement lr/1s-like repeat
F:554-663/Domain: C1r/C1s repeat homology <C1R3>
F:62,105,295,326/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:176,180,186,235/Binding site: zinc (His, His, His, Tyr) #status predicted
F:177/Active site: Glu #status predicted
F:528/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 9.5%; Score 176; DB 2; Length 707;
Best Local Similarity 43.3%; Pred. NO. 3.5e-07;
Matches 45; Conservative 15; Mismatches 36; Indels 8; Gaps 5;

QY 55 NGSITHSPRPHTYPRNTVLVRLVAVEENVWQLTFDERFGLDEPDDICKYDFVEVEE- 113
Db 562 NGSINSPGWPKEYPPNKNKINQVAPQYRISLKFQD---FETEGNDVCKYDFVEVRS 617

Qy 114 --PSDGTILRCWCGSGTVPGKQISKGNQIRIRFVSDYFPPSEPGF 156
Db 618 LTSDSLKLGKFCGS-ELPAVITSOYNNMRIFKSDNTV-SKKGF 659

RESULT 5
A:58788
N:collagen C-endopeptidase (EC 3.4.24.19) precursor, splice form HIS - human
C:Species: Homo sapiens (man)
C:Date: 28-Mar-1998 #sequence_revision 09-Apr-1998 #text_change 18-Jun-1999
C:Accession: A37278; A58788
R:Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitsock, L.M.; Whitters, M.J.; Kriz, R.W.; Hen
Science 242, 1528-1534, 1988
A:Title: Novel regulators of bone formation: molecular clones and activities.
A:Reference number: A37278; MUID:89072730
A:Accession: A37278
A:Molecule type: mRNA
A:Residues: 1-702, 'EKRPALQPPRGRPHQLKFRVQKRNRTPO' <WOZ>
R:Takahara, K.; Lyons, G.E.; Greenspan, D.S.
J. Biol. Chem. 269, 32572-32578, 1994
A:Title: Bone morphogenetic protein-1 and a mammalian tolloid homologue (mtld) are encod
A:Reference number: A58788; MUID:95096114
A:Accession: A58788
A:Molecule type: mRNA
A:Residues: 703-823 <TAK>
A:Cross-references: GB:L35278; NID:g619423; PIDN:AAC41703.1; PID:g619424
C:Genetics:
A:Gene: GDB:BMPI; BMP-1
A:Cross-references: GDB:125203; OMIM:112264
A:Map position: 8p21-8p21
C:Function:
A:Description: catalyzes hydrolysis of the carboxyl-terminal propeptide of collagen type
C:Superfamily: procollagen C-endopeptidase; astacin homology; Clr/Cls repeat homology; H
C:Keywords: alternative splicing; beta-hydroxyasparagine; bone; calcium; duplication; gl
F:1-23/Domain: signal sequence #status predicted <SIG>
F:23-823/Product: procollagen C-endopeptidase splice form HIS #status predicted <MAT>
F:130-321/Domain: astacin homology <AST>
F:322-431/Domain: Clr/Cls repeat homology <CLR1>
F:435-544/Domain: Clr/Cls repeat homology <CLR2>
F:551-587/Domain: EGF homology <EGF>
F:591-700/Domain: Clr/Cls repeat homology <CLR>
F:738-752/Region: histidine-rich
F:91.142,332,363,599/Binding site: carboxylate (Asn) (covalent) #status predicted
F:163-319,185-205,322-348,375-397,435-461,488-510,551-563,559-572,574-587,591-617,644-66
F:213,217,223,272/Binding site: zinc (His, His, His, Tyr) #status predicted
F:214/Active site: Glu #status predicted
F:565/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 9.3%; Score 173; DB 1; Length 823;
Best Local Similarity 35.6%; Pred. No. 7.7e-07;
Matches 52; Conservative 19; Mismatches 49; Indels 26; Gaps 7;

Qy 55 NGSIHSPRPHTYPRNTVLVRLVA-VEENVWQLTFDERFGLDEPDDICKYDFVVEE 113
Db 599 NGSITSPGPKPEYPPNNKNCITQLVAPTYQYRISLQDFDFEG-----NDVCKYDFVEVRS 653

Qy 114 --PSDGTILRCWCGSGTVPGKQISKGNQIRIRFVSDYFPPSEPGFCTHYNVMPQFTEAV 171
Db 654 GLTADSKLHGKFCGS-EKPEVITSOYNNMRIFKSDNTV-SKKGKAHFFSVLEAGDRH 711

Qy 172 S-----PSVL-----PPSAL 181
Db 712 SHLSGLELLICPHALVDVTVPAPPSAL 737

RESULT 6
T08618
Intrinsic factor-B12 receptor CUBILIN precursor - rat
C:Species: Rattus norvegicus (Norway rat)

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T08618
R:Moestrup, S.K.; Kozyraki, R.; Kristiansen, M.; Kaysen, J.H.; Rasmussen, H.H.; Braul
J. Biol. Chem. 273, 5235-5242, 1998
A:Title: The intrinsic factor-vitamin B12 receptor and target of teratogenic antibody
A:Reference number: Z16459; MUID:98148073
A:Accession: T08618
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-3623 <MOE>
A:Cross-references: EMBL:AF022247; NID:g3834379; PIDN:AAC71661.1; PID:g3834380
C:Genetics:
A:Gene: CUBILIN
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: egg yolk; endocytosis; glycoprotein; intestine; kidney; peripheral membra
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-3623/Product: intrinsic factor-B12 receptor CUBILIN #status predicted <MAT>
F:133-164/Domain: EGF homology <EGF1>
F:436-467/Domain: EGF homology <EGF>

Query Match 9.2%; Score 171; DB 2; Length 3623;
Best Local Similarity 24.9%; Pred. No. 7.5e-06;
Matches 92; Conservative 52; Mismatches 138; Indels 88; Gaps 23;

Qy 2 SLFGLLLTALAGORGTQAESNLSSKQFSSNKEQNGVQDPQHERIITVSTNGSIHSP 61
Db 901 SVYNILYTVTFVKSSME-----NRGFTAKFSDDKLECG-----EVLTAFT-GIIEPSP 946

Qy 62 RFPHTYPRNTVLVRLVA-VEENVWQLTFDERFGLDEPDDICKYDFVVEPSPDGTILG 121
Db 947 GHPNYPYRGVNCVTHVV-VQRCQLIRLEPSS-FYLEFHYN--CTNDYLEIYDTAAQTFLG 1002

Qy 122 RWCNGYVPGKQISKGNQIRIRFVSDYFPPSEPGFCHY-----NIVMPQTEAVSPSV 175
Db 1003 RYCGK-SIPPSLTSSNSIKLIFVSDSALAHE-GFSINYEAIDASSVCLDYDTDFNGMLS 1060

Qy 176 LP--PSALPLDL-----LNNAI---TAFSTLED-----LIRYLEPERWQLDLEL 215
Db 1061 SPNFPNNYPSNWEYIRTVGLNQIALHFTDF-TLEDYFGSOCVDVFE-----IRDG 1112

Qy 216 YRPTWQLGKAFVFGKRSRVVDNLNLTIEVRLY-----SCTPRNFVSIRIELKRTDT 268
Db 1113 GYETSPLVG---IY--CGSVLPPTIISHSNLWLFKSDAALTAKGFS----- 1156

Qy 269 IFWPCCLLVKRCGNCACCLHNCNCCQVPSKVTKYHE---VLQLRPKTVGRVGLHKSILT 325
Db 1157 -YWDGS--STGCGGN---LTPPTGLTSPNYPMPYHSSCYWRLEASHG-SPELELEFQ 1208

Qy 326 DVALEHHEEC 335
Db 1209 DFHLEHHPSC 1218

RESULT 7
B58788
Procollagen C-endopeptidase (EC 3.4.24.19) precursor, tolloid-like splice form - huma
N:Alternate names: bone morphogenic protein 1, tolloid-like splice form
C:Species: Homo sapiens (man)
C:Date: 28-Mar-1998 #sequence_revision 09-Apr-1998 #text_change 16-Jul-1999
C:Accession: A37278; B58788
R:Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitsock, L.M.; Whitters, M.J.; Kriz, R.W.;
Science 242, 1528-1534, 1988
A:Title: Novel regulators of bone formation: molecular clones and activities.
A:Reference number: A37278; MUID:89072730
A:Accession: A37278
A:Molecule type: mRNA
A:Residues: 1-702, 'EKRPALQPPRGRPHQLKFRVQKRNRTPO' <WOZ>
A:Cross-references: GB:M22488; NID:g179499; PIDN:AAA51833.1; PID:g179500
R:Takahara, K.; Lyons, G.E.; Greenspan, D.S.
J. Biol. Chem. 269, 32572-32578, 1994
A:Title: Bone morphogenetic protein-1 and a mammalian tolloid homologue (mtld) are en
A:Reference number: A58788; MUID:95096114

Db 193 SGELSSPEYPRYPKLSSTYS-ISLEEGFSVILDFVESDFVETHPETLCPYDFLKIQ-- 249


```
QY 114 PSDGTILGRWCGS-----GTVPGKQ--ISKGNQIRIRFVSDEYFPSE-----PGFC1 158
      :| |||:| | :| |||:| | :| |||:| | :| |||:| | :| |||:| | :| |||:| |
Db 80 -ADKSLGRFCGQLGSPGNPGKKEFNQGNKMLLTHTD--FSNEENGTFMYKGFLA 136
      :| |||:| | :| |||:| | :| |||:| | :| |||:| | :| |||:| | :| |||:| |
QY 159 HYNIV-----MPQ-----166
      :| | |
Db 137 YQAVDLDECASRSKSGEDPQPCQHLCHNXYGVGFCRPGYELQEDRHSCQAECSSE 196
      :| | |
QY 167 -FTEAVS--PSVLPPSALPLDL-LNNAATFSTLEDLRYLEPERWQLDED---LYRPT 219
      :| | | :| | | :| | | :| | | :| | | :| | | :| | | :| | | :| | |
Db 197 LYTEASGVISLEPRSPDPDLRCNYSRVERGLTLHLKFLFEP-----FDIDHQVHCYPY 252
      :| | | :| | | :| | | :| | | :| | | :| | | :| | | :| | |
QY 220 WLLGKAFVFCR-----KSRVVDLNLLEE 244
      :| | | :| | |
Db 253 DQL--QIYANGKNIGFCGKORPPDLDTSSNAVDLLFTFDE 291
      :| | | :| | |

RESULT 15
I54763
Ra-reactive factor (EC 3.4.21.-) 1 precursor - human
N:Alternate names: mannose binding protein-associated serine proteinase 1 (MASP-1)
C:Species: Homo sapiens (man)
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 16-Jun-2000
R;Accession: I54763; JN0883
C;Sato, T.; Endo, Y.; Matsushita, M.; Fujita, T.
Int. Immunol. 6, 665-669, 1994
A:Title: Molecular characterization of a novel serine protease involved in activation of
A:Reference number: I54763; MUID:94289349
A:Accession: I54763
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-699 <SAT>
A:Cross-references: GB:D28593; NID:g790963; PIDN:BAA05928.1; PID:g471128
R;Takada, F.; Takayama, Y.; Hattuse, H.; Kawakami, M.
Biochem. Biophys. Res. Commun. 196, 1003-1009, 1993
A:Title: A new member of the C1s family of complement proteins found in a bactericidal
A:Reference number: JN0883; MUID:94059062
A:Accession: JN0883
A:Molecule type: mRNA
A:Residues: 1-234, 'E', 236-284, 'G', 286-498, 'K', 500-542, 'K', 544-642, 'S', 644-699 <TAK>
A:Cross-references: DBJ:D17525; NID:g439712; PIDN:BAA04477.1; PID:g439713
A:Experimental source: liver
C:Comment: This is a serum bactericidal factor that activates complement C4 and C2 compo
C:Genetics:
A:Gene: GDB:MASP1; GDB:CFARF; CFARF1; PRSS5; MASP
A:Cross-references: GDB:361104; GDB:330954; OMIM:600521
A:Map position: 3q27-3q28
C:Superfamily: complement subcomponent C1r; C1r/C1s repeat homology; complement factor H
C:Keywords: beta-hydroxyasparagine; complement pathway; duplication; glycoprotein; hydro
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-448,449-699/Product: Ra-reactive factor #status predicted <MAT>
F;19-135/Domain: C1r/C1s repeat homology <C1r1>
F;143-181/Domain: EGF homology <EGF>
F;185-294/Domain: C1r/C1s repeat homology <C1r2>
F;301-362/Domain: complement factor H repeat homology <FH1>
F;367-432/Domain: complement factor H repeat homology <FH2>
F;449-691/Domain: trypsin homology <TRY>
F;49,178,407/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;73-91,143-157,153-166,168-181,185-212,242-260,301-349,329-362,367-414,397-432,436-572,
F;159/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F;448-449/cleavage site: Arg-Ile (autolytic) #status predicted
F;490,552,646/Active site: His, Asp, Ser #status predicted

Query Match 7.4%; Score 137.5; DB 1; Length 699;
Best Local Similarity 34.1%; Pred. No. 0.00075;
Matches 43; Conservative 22; Mismatches 40; Indels 21; Gaps 9;

QY 51 TVSTN---GSIHSPRPHYTPNTVLVWRLVAVENWVLTFTDFRFGLEDPEDDICKYD 107
      :| | | :| | | :| | | :| | | :| | | :| | | :| | | :| | |
Db 21 TVELNNMGQIQSPGYPSPDSEVTWN-ITVPDGRFKLYF-MHFNLE--SSYLCEYD 76
      :| | | :| | | :| | | :| | | :| | | :| | | :| | | :| | |
QY 108 FVEVEPSDGTILGRWCGSGT-----VPGKQI--SKGNQIRIRFVSDEYFPSE---PGFC 157
```

```
Db 77 YKVE--TEDQVLATFCGRETITDTEQTPGQEVVLSGSEFNSITFRSD--FSNEERFTGFD 132
      :| | | :| | | :| | | :| | | :| | | :| | | :| | | :| | |
QY 158 IHNIV 163
      :| | |
Db 133 AHYMAV 138
      :| | |
```

Search completed: June 29, 2001, 17:09:52
Job time: 184 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2001, 17:08:48 ; Search time 12.01 Seconds
(without alignments)
984.025 Million cell updates/sec

Title: US-09-457-066-2
Perfect score: 1858
Sequence: 1 MSLLFGLLLTSALAQORQT.....DVALEHHECDVCVRGSGTG 345

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 179.5 | 9.7 | 928 | 1 NRPI_XENLA | P28824 xenopus lae |
| 2 | 176 | 9.5 | 707 | 1 BMP1_XENLA | P98070 xenopus lae |
| 3 | 173.5 | 9.3 | 616 | 1 SPAN_STRPU | P98068 strongyloce |
| 4 | 169 | 9.1 | 986 | 1 BMP1_HUMAN | P13497 homo sapien |
| 5 | 169 | 9.1 | 991 | 1 BMP1_MOUSE | P98063 mus musculu |
| 6 | 163.5 | 8.8 | 922 | 1 NRPI_RAT | Q9qwj9 rattus norv |
| 7 | 163.5 | 8.8 | 931 | 1 NRP2_HUMAN | O60462 homo sapien |
| 8 | 163 | 8.8 | 925 | 1 NRP2_RAT | O35276 rattus norv |
| 9 | 160.5 | 8.6 | 923 | 1 NRP1_MOUSE | P97333 mus musculu |
| 10 | 159.5 | 8.6 | 931 | 1 NRP2_MOUSE | O35375 mus musculu |
| 11 | 158.5 | 8.5 | 923 | 1 NRP1_HUMAN | O44786 homo sapien |
| 12 | 157 | 8.4 | 914 | 1 NRPI_CHICK | P79795 gallus gall |
| 13 | 154 | 8.3 | 449 | 1 PCOL_HUMAN | Q15113 homo sapien |
| 14 | 148.5 | 8.0 | 1057 | 1 TLD_DROME | P25723 drosophila |
| 15 | 143.5 | 7.7 | 686 | 1 MAS2_HUMAN | O00187 homo sapien |
| 16 | 139.5 | 7.5 | 597 | 1 BP10_PARLI | P42674 paracentrot |
| 17 | 138.5 | 7.5 | 704 | 1 CRAR_MOUSE | P98064 mus musculu |
| 18 | 138 | 7.4 | 639 | 1 BMP1_STRPU | P98069 strongyloce |
| 19 | 138 | 7.4 | 705 | 1 C1R_HUMAN | P00736 homo sapien |
| 20 | 137.5 | 7.4 | 699 | 1 CRAR_HUMAN | P48740 h complemen |
| 21 | 132 | 7.1 | 419 | 1 VEGC_HUMAN | P49767 homo sapien |
| 22 | 128 | 6.9 | 468 | 1 PCOL_MOUSE | Q61398 mus musculu |
| 23 | 128 | 6.9 | 468 | 1 PCOL_RAT | O08628 rattus norv |
| 24 | 125 | 6.7 | 213 | 1 PDGA_XENLA | P34007 oryctolagus |
| 25 | 125 | 6.7 | 415 | 1 VEGC_MOUSE | P97953 mus musculu |
| 26 | 123.5 | 6.6 | 277 | 1 TSG6_HUMAN | P98066 homo sapien |
| 27 | 120.5 | 6.5 | 245 | 1 PDGB_FELCA | P12919 felis silve |
| 28 | 118.5 | 6.4 | 276 | 1 TSG6_RABIT | P98065 oryctolagus |
| 29 | 114.5 | 6.2 | 148 | 1 VEGH_ORFN7 | P52585 orf virus (|
| 30 | 114.5 | 6.2 | 241 | 1 PDGB_HUMAN | P01127 homo sapien |
| 31 | 112.5 | 6.1 | 226 | 1 PDGA_XENLA | P13698 xenopus lae |
| 32 | 111.5 | 6.0 | 514 | 1 UVS2_XENLA | P42664 xenopus lae |
| 33 | 110 | 5.9 | 211 | 1 PDGA_MOUSE | P20033 mus musculu |

RESULT 1

| ID | NRPI_XENLA | STANDARD; | PRT; | 928 AA. |
|----|--|-----------|------|---------|
| AC | P28824: | | | |
| DT | 01-DEC-1992 (Rel. 24, Created) | | | |
| DT | 01-DEC-1992 (Rel. 24, Last sequence update) | | | |
| DT | 01-OCT-2000 (Rel. 40, Last annotation update) | | | |
| DE | NEUROFILIN-1 PRECURSOR (A5 PROTEIN) (A5 ANTIGEN). | | | |
| OS | Xenopus laevis (African clawed frog). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; | | | |
| OC | Xenopodinae; Xenopus. | | | |
| OX | NCBI_TaxID=8355; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Brain; | | | |
| RA | MEDLINE=91337458; PubMed=1908252; | | | |
| TX | Tagaki S., Hirata T., Agata K., Mochii M., Eguchi G., Fujisawa H.; | | | |
| RT | "The A5 antigen, a candidate for the neuronal recognition molecule, | | | |
| RT | has homologues to complement components and coagulation factors."; | | | |
| RL | Neuron 7:295-307(1991). | | | |
| CC | !- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE | | | |
| CC | CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF | | | |
| CC | CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS OUTSIDE THE NERVOUS | | | |
| CC | SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS (BY | | | |
| CC | SIMILARITY). PRESUMED TO BE INVOLVED IN THE NEURONAL RECOGNITION | | | |
| CC | BETWEEN THE OPTIC NERVE FIBERS AND THE VISUAL CENTERS. | | | |
| CC | !- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. | | | |
| CC | !- TISSUE SPECIFICITY: RETINAL GANGLION CELLS AND VISUAL CENTER | | | |
| CC | NEURONS. | | | |
| CC | !- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY. | | | |
| CC | !- SIMILARITY: CONTAINS 2 CUB DOMAINS. | | | |
| CC | !- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS. | | | |
| CC | !- SIMILARITY: CONTAINS 1 MAM DOMAIN. | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL Outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | |
| CC | or send an email to license@isb-sib.ch). | | | |
| CC | ----- | | | |
| DR | EMBL; D10467; BAA01260.1; - | | | |
| DR | InterPro; IPR000421; - | | | |
| DR | InterPro; IPR000859; - | | | |
| DR | InterPro; IPR000998; - | | | |
| DR | Pfam; PF00431; CUB; 2. | | | |
| DR | Pfam; PF00754; F5_F8_type_C; 2. | | | |
| DR | Pfam; PF00629; MAM; 1. | | | |
| DR | PRINTS; PRO0020; MAMDOMAIN. | | | |
| DR | PROSITE; PS00740; MAM_1; 1. | | | |
| DR | PROSITE; PS01180; CUB; 2. | | | |
| DR | PROSITE; PS01285; FA58C_1; 2. | | | |
| DR | PROSITE; PS01286; FA58C_2; 2. | | | |
| DR | PROSITE; PS00600; MAM_2; 1. | | | |

| | | | | | | |
|----|-------|-----|-----|---|------------|--------------------|
| 34 | 109 | 5.9 | 204 | 1 | PDGA_RAT | P28576 rattus norv |
| 35 | 108.5 | 5.8 | 164 | 1 | VEGF_CAVPO | P26617 cavia porce |
| 36 | 108.5 | 5.8 | 226 | 1 | TSIS_SMSAV | P01128 simian sarc |
| 37 | 105.5 | 5.7 | 188 | 1 | VEGB_HUMAN | P49765 homo sapien |
| 38 | 104.5 | 5.6 | 215 | 1 | VEGF_HUMAN | P15692 homo sapien |
| 39 | 104 | 5.6 | 211 | 1 | PDGA_HUMAN | P04085 homo sapien |
| 40 | 102.5 | 5.5 | 133 | 1 | VEGH_ORFN2 | P52584 orf virus (|
| 41 | 102.5 | 5.5 | 190 | 1 | VEGF_PIG | P49151 sus scrofa |
| 42 | 101.5 | 5.5 | 146 | 1 | VEGF_SHEEP | P50412 ovus aries |
| 43 | 101.5 | 5.5 | 190 | 1 | VEGF_BOVIN | P15691 bos taurus |
| 44 | 101.5 | 5.5 | 225 | 1 | PDGB_RAT | Q05028 rattus norv |
| 45 | 98.5 | 5.3 | 241 | 1 | PDGB_SHEEP | Q95229 ovus aries |

ALIGNMENTS

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL: L12249; AAA16313.1; -
DR HSP; P00736; LAPO.
DR MEROPS; M12.005;
DR InterPro; IPR000130; -
DR InterPro; IPR000152; -
DR InterPro; IPR000561; -
DR InterPro; IPR000859; -
DR InterPro; IPR001506; -
DR InterPro; IPR001881; -
DR Pfam; PF01400; Astacin; 1.
DR Pfam; PF00431; CUB; 3.
DR Pfam; PF00008; EGF; 1.
DR PRINTS; PRO0480; ASTACIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS01180; CUB; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
KW Growth factor; Cytokine; Repeat; Bone; Cartilage; Hydrolase; Protease;
KW Metalloprotease; EGF-like domain; Zinc; Calcium; Signal;
KW Glycoprotein. 1 ? POTENTIAL.
FT SIGNAL 1 ? POTENTIAL.
FT PROPEP 83
FT CHAIN 84 707 BONE MORPHOGENETIC PROTEIN 1.
FT DOMAIN 84 284 METALLOPROTEASE.
FT FT 285 397 CUB.
FT DOMAIN 398 509 CUB.
FT DOMAIN 510 551 EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 554 666 CUB.
FT METAL 176 176 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 177 177 EGF-LIKE.
FT METAL 180 180 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 514 526 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 522 535 BY SIMILARITY.
FT DISULFID 537 550 BY SIMILARITY.
FT CARBOHYD 62 62 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 326 326 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 562 562 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE 707 AA; 80673 MW; 1B6980D716DC9B8D CRC64;

Query Match 9.5%; Score 176; DB 1; Length 707;
Best Local Similarity 43.3%; Pred. No. 2e-07;
Matches 45; Conservative 15; Mismatches 36; Indels 8; Gaps 5;

QY 55 NGSIHSPRPHTYPRNTVLVWRLVAVEENVWIGLTDEREGLEDPEDDICKYDFVEE- 113
| | | | | : | | : | | | : | | | : | | | | |
DB 562 NGSINSPGWPKEYPPNKNKINQIWLVAPTQ-YRISLKFDQ---FTEGNDVKCYDFVEVRSG 617
| | | | | : | | : | | | : | | | : | | | | |
QY 114 -PSDGITLGRWCSSGTVPGKQISKGNQIRFRFVSDEYFPSEPFG 156
| | | | | : | | : | | | : | | | : | | | | |
DB 618 LTSDSKLHFKCGS-ELPAVITTSQYNNMRFKSDNTV-SKRGF 659
| | | | | : | | : | | | : | | | : | | | | |

RESULT 3
SPAN_STRPU ID SPAN_STRPU STANDARD; PRT; 616 AA.
AC P98068;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)

| | | |
|------------|--|-----------------------|
| 349 | GETSPNYSNVEDNTACVYEIEGPGVST-IELTF---LDMEIETETLCRYDAVEVRKDD | 404 |
| 116 | DGTLIRGCGSGTVPKQISKGNQIRIRFVSDYEPSP--EPGFCIHYNVMQPFTEAVSP | 173 |
| 405 | INSIGKEKFCGN-TLPPVQIISNNQMVSTSD---PSITRRGFKATYVIII-QTTTVEST | 459 |
| 174 | SVL---PPSALPLDLLNNAITAFSTLEDLIRYLEPERWQOLDLEDLYRPTWQLLKAFYFG | 230 |
| 460 | TTLQTTTPPTSTTLQTTNPSTTTLQT---TNPSTTTTLQTTD--TPVIGSCGCTEV-G | 509 |
| 231 | RKSRVVDLMLLTVEVRLYSCTPRNFVSISR-----ELKKTDRIF-----W | 271 |
| 510 | VEGRVASPNY-----PNDYDNLQCDYIEVDGRRVELIFEDFGLEDFTTCRW | 558 |
| 272 | PGCLL-----VKRCG 281 | |
| 559 | DSLIMNLNGIKRVKMKCG 577 | |
| RESULT | 4 | |
| BMP1_HUMAN | | |
| ID | BMP1_HUMAN | STANDARD: PRT: 986 AA |
| AC | P13497; Q13292; Q99421; Q99422; Q99423; Q14874; | |
| DT | 01-JAN-1990 (Rel. 13, Created) | |
| DT | 01-OCT-2000 (Rel. 40, Last sequence update) | |
| DT | 01-OCT-2000 (Rel. 40, Last annotation update) | |
| DE | BONE MORPHOGENETIC PROTEIN 1 PRECURSOR (EC 3.4.24.19) (BMP-1) | |
| DE | (PROCOLLAGEN C-PROTEINASE) (PCP) (MAMMALIAN TOLLOID PROTEIN) (MTLD). | |
| GN | BMP1. | |
| OS | Homo sapiens (Human). | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| OX | NCBI_TaxID=9606; | |
| ON | [1] | |
| RP | SEQUENCE FROM N.A. (ISOFORM BMP1-3). | |
| RP | TISSUE=SkIn; | |
| RX | PubMed=8643539; | |
| RX | Li S.W., Sieron A.L., Fertala A., Hojima Y., Arnold W.V., | |
| RA | Prockop D.J.; | |
| RT | "The C-proteinase that processes procollagens to fibrillar collagens | |
| RT | is identical to the protein previously identified as bone morphogenic | |
| RT | protein-1."; | |
| RT | Proc. Natl. Acad. Sci. U.S.A. 93:5127-5130(1996). | |
| RN | [2] | |
| RN | SEQUENCE FROM N.A. (ISOFORM BMP1-1). | |
| RP | MEDLINE=89072730; PubMed=3201241. | |
| RX | Wozney J.M., Rosen V., Celeste A.J., Mitsock L.M., Whitters M.J., | |
| RA | Kriz R.W., Hewick R.M., Wang E.A.; | |
| RT | "Novel regulators of bone formation: molecular clones and | |
| RT | activities."; | |
| RT | Science 242:1528-1534(1988). | |
| RN | [3] | |
| RP | SEQUENCE FROM N.A. (ISOFORMS BMP1-4; BMP1-5 AND BMP1-6). | |
| RP | TISSUE=Placenta; | |
| RX | PubMed=9500680; | |
| RA | Janitz M., Heiser V., Boettcher U., Landt O., Lauster R.; | |
| RT | "Three alternatively spliced variants of the gene coding for the human | |
| RT | bone morphogenetic protein-1;" | |
| RT | J. Mol. Med. 76:141-146(1998). | |
| RN | [4] | |
| RP | SEQUENCE OF 703-986 FROM N.A. (ISOFORM BMP1-3). | |
| RP | TISSUE=Placenta; | |
| RX | MEDLINE=95096114; PubMed=7798260; | |
| RA | Takahara K., Lyons G.E., Greenspan D.S.; | |
| RT | "Bone morphogenetic protein-1 and a mammalian tolloid homologue (mtld) | |
| RT | are encoded by alternatively spliced transcripts which are | |
| RT | differentially expressed in some tissues."; | |
| RT | J. Biol. Chem. 269:32572-32578(1994). | |
| CC | - I- FUNCTION: CLEAVES THE C-TERMINAL PROPEPTIDES OF PROCOLLAGEN I, II | |
| CC | AND II. INDUCES CARTILAGE AND BONE FORMATION. | |
| CC | - I- CATALYTIC ACTIVITY: CLEAVAGE OF THE C-TERMINAL PROPEPTIDE AT | |
| CC | ALA-I-ASP IN TYPE I AND II PROCOLLAGENS AND AT ARG-I-ASP IN TYPE | |
| CC | II. | |

| | | | | |
|-----|----------|--|-----|-------------------------------------|
| FT | DISULFID | 28 | 55 | BY SIMILARITY. |
| FT | DISULFID | 83 | 105 | BY SIMILARITY. |
| FT | DISULFID | 149 | 175 | BY SIMILARITY. |
| FT | DISULFID | 208 | 230 | BY SIMILARITY. |
| FT | DISULFID | 277 | 427 | BY SIMILARITY. |
| FT | DISULFID | 434 | 592 | BY SIMILARITY. |
| FT | CARBOHYD | 152 | 152 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 157 | 157 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 629 | 629 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 833 | 833 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 834 | 834 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| SEQ | SEQUENCE | 925 AA; 103896 MW; 3BP62903F644851C CRC64; | | |

Query Match 8.8%; Score 163; DB 1; Length 925;
 Best Local Similarity 33.8%; Pred. No. 3.7e-06;
 Matches 45; Conservative 20; Mismatches 60; Indels 8; Gaps

| | | | |
|----|-----|---|--|
| QY | 32 | FSSNKEONGVQDQOHERLIITVSTNGSTHSRPFHTYPRNTVLVWRLVAVEENVWIOITD 919 | |
| DB | 15 | FSGHKVRN-QODPCGGRLNSKDAGYITSPGYQDYPHQNCSEWVYAPEPNQIKVLNFN 73 | |
| QY | 92 | ERFLEQDEDDICKYDEVEVE--EPSDGTILGRWCWGSGTVPGRQISGKNGOIRIRFVSDEY 149 | |
| DB | 74 | PHPELEKHD--CKYDFIETRDGDSADLLGRKCGN-IAPPIISGGSVLYIKFTSD-Y 128 | |
| QY | 150 | FPSPGPGCIHNYI 162 | |
| DB | 129 | ARQAGGFSRLYEI 141 | |

RESULT 9

| | | | | |
|----|--|-----------|------|---------|
| ID | NRPI_MOUSE | STANDARD; | PRT; | 923 AA. |
| AC | P97333; | | | |
| DT | 01-NOV-1997 (Rel. 35, Created) | | | |
| DT | 01-NOV-1997 (Rel. 35, Last sequence update) | | | |
| DT | 01-OCT-2000 (Rel. 40, Last annotation update) | | | |
| DE | NEUROFILIN-1 PRECURSOR (A5 PROTEIN). | | | |
| GN | NRPI OR NRP. | | | |
| OS | Mus musculus (Mouse). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| OX | NCBI_TaxID=10090; | | | |
| RP | {1} | | | |
| RN | SEQUENCE FROM N.A. | | | |
| RC | SPRAIN=BALB/C; TISSUE=Embryonic brain; | | | |
| RA | MEDLINE=96353149; PubMed=8748368; | | | |
| XX | Kawakami A., Kitsuakawa T., Takagi S., Fujisawa H.; | | | |
| RT | "Developmentally regulated expression of a cell surface protein, | | | |
| RT | neuropilin, in the mouse nervous system."; | | | |
| RL | J. Neurobiol. 29:1-17(1996). | | | |
| CC | - FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE | | | |
| CC | CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF | | | |
| CC | CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS OUTSIDE THE NERVOUS | | | |
| CC | SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS. I | | | |
| CC | BINDS TO SEMAPHORIN 3A, THE PLGF-2 ISOFORM OF PGF, THE VEGF-165 | | | |
| CC | ISOFORM OF VEGF AND VEGF-B. COEXPRESSION WITH KDR RESULTS IN | | | |
| CC | INCREASED VEGF-165 BINDING TO KDR AS WELL AS INCREASED CHEMOTAXIS | | | |
| CC | IT MAY REGULATE VEGF-INDUCED ANGIOGENESIS (BY SIMILARITY). | | | |
| CC | - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. | | | |
| CC | - TISSUE SPECIFICITY: NERVOUS SYSTEM. | | | |
| CC | - SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY. | | | |
| CC | - SIMILARITY: CONTAINS 2 CUB DOMAINS. | | | |
| CC | - SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS. | | | |
| CC | - SIMILARITY: CONTAINS 1 NAM DOMAIN. | | | |

This SWISS-PROT entry is copyright. It is produced through a collabora-
 between the Swiss Institute of Bioinformatics and the EMBL outstat-
 the European Bioinformatics Institute. There are no restrictions on
 use by non-profit institutions as long as its content is in no
 modified and this statement is not removed. Usage by and for comm-
 entities requires a license agreement (See <http://www.isb-sib.ch/annou>)

```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D50086; BAA08789.1; -.
CC MGD; MGI:106206; Nrp.
CC InterPro; IPR000421; -.
CC InterPro; IPR000859; -.
CC InterPro; IPR000998; -.
CC Pfam; PF00754; F5_F8_type_C; 2.
CC Pfam; PF00754; F5_F8_type_C; 2.
CC Pfam; PF00629; MAM; 1.
CC PROSITE; PS01180; CUB; 2.
CC PROSITE; PS01285; FA58C_1; 2.
CC PROSITE; PS01286; FA58C_2; 2.
CC PROSITE; PS00740; MAM_1; 1.
CC PROSITE; PS00600; MAM_2; 1.
CC Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor.
FT SIGNAL; 1 21
FT CHAIN; 22 923
FT DOMAIN; 22 856
FT TRANSMEM; 857 879
FT DOMAIN; 880 923
FT DOMAIN; 27 141
FT DOMAIN; 147 265
FT DOMAIN; 275 424
FT DOMAIN; 431 583
FT DOMAIN; 645 811
FT DISULFID; 27 54
FT DISULFID; 82 104
FT DISULFID; 147 173
FT DISULFID; 206 228
FT DISULFID; 275 424
FT DISULFID; 431 583
FT CARBOHYD; 150 150
FT CARBOHYD; 261 261
FT CARBOHYD; 300 300
FT CARBOHYD; 522 522
FT CARBOHYD; 842 842
FT SSEQUENCE; 923 AA; 103020 MW; 0644B8A10796808 CRC64;
SQ
Query Match 8.68; Score 160.5; DB 1; Length 923;
Best Local Similarity 36.58; Pred. No. 6.1e-06;
Matches 42; Conservative 16; Mismatches 50; Indels 7; Gaps 4;
QY 50 ITVSTNGSIHSPRPHYPRTVLVRLVAVENNVQLTDERFGLGLEDPPDICKYDFV 109
Db 31 IKIENPCYLSPGYPHSPHSEKCEWLQAPQYRIINPHFDLEDRD---CKYDV 87
QY 110 EV--EPPSDGTILGRWCGSGTVPGKQISGNQIRIRFVSDYFPPSEPGFCIHNI 162
Db 88 EVIDGENEGRLMGKFCGK-IAPSPVSSGPPFLFKFVSD-YETHGAGFSIRYEI 140
RESULT 10
NRP2_MOUSE STANDARD; PRT; 931 AA.
AC O35375; O35373; O35374; O35376; O35377; O35378;
AT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROFILIN-2 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165
DE RECEPTOR 2).
GN NRP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX STRAIN=BALB/C;
MEDLINE=97470888; PubMed=9331348;
RA Chen H., Chedotal A., He Z.-G., Goodman C.S., Tessier-Lavigne M.;
RT "Neurofilin-2, a novel member of the neurofilin family, is a high
```

```
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RT III.";
RL Neuron 19:547-559(1997).
CC -!- FUNCTION: HIGH AFFINITY RECEPTOR FOR SEMAPHORINS 3C, 3F, VEGF-165
CC AND VEGF-145 ISOFORMS OF VEGF, AND THE PLGF-2 ISOFORM OF PGF.
CC -!- SUBUNIT: NEUROFILIN-2 PROBABLY FORMS AN HETEROMERIC COMPLEX WITH
CC NEUROFILIN-1 IN ORDER TO BE A FUNCTIONAL SEMAPHORIN E RECEPTOR.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- ALTERNATIVE PRODUCTS: 6 ISOFORMS; A22 (SHOWN HERE), A0, A5, A17,
CC B0 AND B5; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN DEVELOPING CNS, PNS AND IN SOME
CC NONNEURAL TISSUES INCLUDING LIMB BUDS, DEVELOPING BONES, MUSCLES,
CC INTESTINAL EPITHELIUM, KIDNEY, LUNG AND SUBMANDIBULAR GLAND.
CC -!- DEVELOPMENTAL STAGE: THE EXPRESSION PATTERN IS VERY DYNAMIC AND
CC IS DEVELOPMENTALLY REGULATED.
CC -!- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF022856; AAC53379.1; -.
CC EMBL; AF022854; AAC53377.1; -.
CC EMBL; AF022855; AAC53378.1; -.
CC EMBL; AF022857; AAC53380.1; -.
CC EMBL; AF022858; AAC53381.1; -.
CC EMBL; AF022861; AAC53382.1; -.
CC MGD; MGI:1100492; Nrp2.
CC InterPro; IPR000421; -.
CC InterPro; IPR000859; -.
CC InterPro; IPR000998; -.
CC Pfam; PF00431; CUB; 2.
CC Pfam; PF00754; F5_F8_type_C; 2.
CC PROSITE; PS00629; MAM; 1.
CC PROSITE; PS01180; CUB; 2.
CC PROSITE; PS01285; FA58C_1; 2.
CC PROSITE; PS01286; FA58C_2; 2.
CC PROSITE; PS00600; MAM_2; 1.
CC Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
KW Alternative splicing.
FT SIGNAL; 1 20
FT CHAIN; 21 931
FT DOMAIN; 21 864
FT TRANSMEM; 865 889
FT DOMAIN; 890 931
FT DOMAIN; 28 142
FT DOMAIN; 149 267
FT DOMAIN; 277 427
FT DOMAIN; 434 592
FT DOMAIN; 642 802
FT DOMAIN; 838 845
FT DISULFID; 28 55
FT DISULFID; 83 105
FT DISULFID; 149 175
FT DISULFID; 208 230
FT DISULFID; 277 427
FT DISULFID; 434 592
FT CARBOHYD; 152 152
FT CARBOHYD; 157 157
FT CARBOHYD; 629 629
FT CARBOHYD; 839 839
FT CARBOHYD; 809 813
FT VARSPIC; 809 830
FT VARSPIC; 814 830
FT VARSPIC; 810 931
FT -----
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MISSING (IN ISOFORM A17).
FT MISSING (IN ISOFORM A0).
FT MISSING (IN ISOFORM A5).
FT SSGKESWLYTLDPILITITIAMSSSLGVLGATCAGLLLYCT
```

```
FT  CTSYGLSSRSCTTLENYFELYDGLKHKVKINHQCSEA
FT  -> GTPPGTPTVDTPVQIPAYWYVMAAGGAVLVLAS
FT  WVALVHYHFRFAAKTDSITYKTSYTNAPLAVEPT
FT  LTIKLEQRGSHC (IN ISOFORM B0).
FT  VDIPETHGEGVEDEIDDEYEDGWSNSSTSGAGDPSSGK
FT  EKSWLYTLDPIILIIAMSSLGVLGACAGLLIYCTCSYS
FT  GLSSRSCTTLENYFELYDGLKHKVKINHQCSEA -> G
FT  GTPPGTPTVDTPVQIPAYWYVMAAGGAVLVLASVVL
FT  ALVLYHFRFAAKTDSITYKTSYTNAPLAVEPTITI
FT  KLEQRGSHC (IN ISOFORM B5).
FT  G -> I (IN AAC53380 AND AAC53381).
FT  CONFLICT 786 786
FT  SEQUENCE 931 AA; 104558 MW; 76F2443F411D2F63 CRC64;
SQ

Query Match 8.6%; Score 159.5; DB 1; Length 931;
Best Local Similarity 33.1%; Pred. No. 7.5e-06;
Matches 44; Conservative 17; Mismatches 65; Indels 7; Gaps 4;

QY 32 FSSNKEGVODPOHERIITVSTNGSIHSPRPHTYPRNTVLVRLVAEENWVQLTPTD 91
   : | | | | | : : : | | | : | | | : | | | : | | | : | | | :
Db 14 YFSGHEVRSQDPPCGGRPNKSDAGIYITSPGYQDYPHQNCWIVYAPEPNQKVLNEN 73
   : | | | | | : : : | | | : | | | : | | | : | | | : | | | :
QY 92 ERFGLEPDDICKYDFEVE--EPSDGTILGRWCGSGTVPQKQISKGNIQIRFVSDRY 149
   : | | : | | | | | : : : | | | : | | | : | | | : | | | :
Db 74 PHFEIKHD---CRYDFEIRDGDSESADLLGKHGCGN-IAPPTIISGSLVYIKETSD-Y 128
   : | | | | | : : : | | | : | | | : | | | : | | | : | | | :
QY 150 FPSEPGFCIHVNI 162
   : | | | | | : : : | | | : | | | : | | | : | | | : | | | :
Db 129 ARQAGGFSRVEI 141
   : | | | | | : : : | | | : | | | : | | | : | | | : | | | :

RESULT 11
NRPI_HUMAN STANDARD; PRT; 923 AA.
ID O14786; O60461;
DT 30-MAY-2000 (Rel. 39, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROPILIN-1 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165
DE RECEPTOR).
GN NRPI OR NRP OR VEGF165R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM).
RX MEDLINE=97433084; PubMed=9288753;
RA He Z., Tessier-Lavigne M.;
RT "Neuropilin is a receptor for the axonal chemorepellent semaphorin
RT III".
RL Cell 90:739-751(1997).
RN [2]
RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM), AND SEQUENCE OF 22-39.
RC TISSUE=Breast;
RX MEDLINE=98188099; PubMed=9529250;
RA Soker S., Takashima S., Mao H.-Q., Neufeld G., Klagsbrun M.;
RT "Neuropilin-1 is expressed by endothelial and tumor cells as an
RT isoform-specific receptor for vascular endothelial growth factor.";
RL Cell 92:735-745(1998).
RN [3]
RP SEQUENCE FROM N.A. (SOLUBLE/SNRP1 ISOFORM), AND SEQUENCE OF 22-31.
RC TISSUE=Prostatic adenocarcinoma;
RX MEDLINE=20183929; PubMed=10688880;
RA Gagnon M.L., Bieleberg D.R., Gerechtman Z., Miao H.-Q., Takashima S.,
RA Soker S., Klagsbrun M.;
RT "Identification of a natural soluble neuropilin-1 that binds vascular
RT endothelial growth factor: In vivo expression and antitumor
RT activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2573-2578(2000).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=20309748; PubMed=10748121;
```

```
RA Gluzman-Poltorak Z., Cohen T., Herzog Y., Neufeld G.;
RT "Neuropilin-2 and neuropilin-1 are receptors for the 165-amino acid
RT form of vascular endothelial growth factor (VEGF) and of placenta
RT growth factor-2, but only neuropilin-2 functions as a receptor for
RT the 145-amino acid form of VEGF.";
RL J. Biol. Chem. 275:18040-18045(2000).
CC -!- FUNCTION: THE MEMBRANE-BOUND ISOFORM IS A RECEPTOR INVOLVED IN THE
CC DEVELOPMENT OF THE CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE
CC FORMATION OF CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS
CC OUTSIDE THE NERVOUS SYSTEM. IT MEDIATES THE CHEMOREPULSANT
CC ACTIVITY OF SEMAPHORINS. IT BINDS TO SEMAPHORIN 3A, THE PLGF-2
CC ISOFORM OF PGF. THE VEGF-165 ISOFORM OF VEGF AND VEGF-B.
CC COEXPRESSION WITH KDR RESULTS IN INCREASED VEGF-165 BINDING TO KDR
CC AS WELL AS INCREASED CHEMOTAXIS. IT MAY REGULATE VEGF-INDUCED
CC ANGIOGENESIS.
CC -!- FUNCTION: THE SOLUBLE/SNRP1 ISOFORM BINDS VEGF-165 AND APPEARS
CC TO INHIBIT ITS BINDING TO CELLS. IT MAY ALSO INDUCE APOPTOSIS BY
CC SEQUESTERING VEGF-165. MAY BIND AS WELL VARIOUS MEMBERS OF THE
CC SEMAPHORIN FAMILY. ITS EXPRESSION HAS AN AVERSE EFFECT ON BLOOD
CC VESSEL NUMBER AND INTEGRITY.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. THE SOLUBLE/SNRP1
CC ISOFORM IS SECRETED.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; MEMBRANE-BOUND (SHOWN HERE)
CC AND SOLUBLE/SNRP1; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: THE EXPRESSION OF MEMBRANE-BOUND AND
CC SOLUBLE/SNRP1 ISOFORMS DOES NOT SEEM TO OVERLAP. THE MEMBRANE-
CC BOUND ISOFORM IS EXPRESSED BY THE BLOOD VESSELS OF DIFFERENT
CC TISSUES. IN THE DEVELOPING EMBRYO IT IS FOUND PREDOMINANTLY IN THE
CC NERVOUS SYSTEM. IN ADULT TISSUES, IT IS HIGHLY EXPRESSED IN HEART
CC AND PLACENTA; MODERATELY IN LUNG, LIVER, SKELETAL MUSCLE, KIDNEY
CC AND PANCREAS; AND LOW IN ADULT BRAIN. THE SOLUBLE/SNRP1 ISOFORM IS
CC FOUND IN LIVER HEPATOCYTES, KIDNEY DISTAL AND PROXIMAL TUBULES.
CC -!- SIMILARITY: BELONGS TO THE NEUROPILIN FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF018956; AAC51759.1; -
CC EMBL; AF016050; AAC12921.1; -
CC EMBL; AF145712; AAF44344.1; -
CC MIM; 602069; -
CC InterPro; IPR000421; -
CC InterPro; IPR000859; -
CC InterPro; IPR000998; -
CC Pfam; PF00431; CUB; 2.
CC Pfam; PF00754; F5_F8_type_C; 2.
CC Pfam; PF00629; MAM; 1.
CC PRINTS; PR00020; MAMDOMAIN.
CC PROSITE; PS01180; CUB; 2.
CC PROSITE; PS01285; FA58C_1; 2.
CC PROSITE; PS01286; FA58C_2; 2.
CC PROSITE; PS00740; MAM_1; 1.
CC PROSITE; PS00600; MAM_2; 1.
CC Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
KW Alternative splicing.
FT SIGNAL 1 21
FT CHAIN 22 923
FT DOMAIN 22 856
FT TRANSMEM 857 879
FT DOMAIN 880 923
FT DOMAIN 27 141
FT DOMAIN 147 265
FT DOMAIN 275 424
FT DOMAIN 431 583
FT DOMAIN 645 811
FT NEUROPILIN-1.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT CUB 1.
FT CUB 2.
FT F5/8 TYPE C 1.
FT F5/8 TYPE C 2.
FT MAM.
```

```
FT DISULFID 27 54 PROBABLE.
FT DISULFID 82 104 PROBABLE.
FT DISULFID 147 173 PROBABLE.
FT DISULFID 206 228 PROBABLE.
FT DISULFID 275 424 BY SIMILARITY.
FT DISULFID 431 583 BY SIMILARITY.
FT CARBOHYD 150 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 842 842 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VASPLIC 642 844 EFP->G1K (IN SOLUBLE/SNRPI ISOFORM).
FT VASPLIC 645 923 MISSING (IN SOLUBLE/SNRPI ISOFORM).
FT CONFLICT 26 26 K -> E (IN REF. 1).
FT CONFLICT 749 749 D -> H (IN REF. 2).
FT CONFLICT 855 855 E -> D (IN REF. 2).
SQ SEQUENCE 923 AA; 103120 MW; ADEADC4A849E5D57 CRC64;
```

Query Match 8.5%; Score 158.5; DB 1; Length 923;
Best Local Similarity 35.7%; Pred. No. 9e-06;
Matches 41; Conservative 18; Mismatches 49; Indels 7; Gaps 4;

```
Qy 50 IVTVSTNGSHSPRPHYPRNTVLVRLVAEENWVLIQTFDERFGLGLEDPEDDICKYDFV 109
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 31 IKIESPGVLTSPGYSPHSEKCEWLIQAPPYQIRIMFNPHFDLED RD---CKYDYV 87
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 110 EV--EPPSDGTILGRWCGSGTVPGKQISGNQIRIRFVSDYFPPSPGFCIHNYI 162
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 88 EYFDGNGENHGRGFCGK-IAPPPVSSGPFLEIKFVSD-YETHGAGFSIRVEI 140
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 12
NRPI_CHICK
ID NRPI_CHICK STANDARD; PRT; 914 AA.
AC P79795;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROFILIN-1 PRECURSOR (A5 PROTEIN).
GN NRPI OR NRP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEHORN; TISSUE=Embryonic brain;
RX MEDLINE=95324761; PubMed=7601310;
RA Takagi S., Kasuya Y., Shimizu M., Matsuura T., Tsuboi M., Kawakami A.,
RA Fujisawa H.,
RT "Expression of a cell adhesion molecule, neuropilin, in the
RT developing chick nervous system.";
RL Dev. Biol. 170:207-222(1995).
CC -!- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE
CC CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF
CC CERTAIN NEURONAL CIRCUITS AND IN ORGANOCENESIS OUTSIDE THE NERVOUS
CC SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS (BY
CC SIMILARITY). SEEMS TO HAVE CALCIUM-INDEPENDENT CELL ADHESION
CC PROPERTIES.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: DEVELOPING NERVOUS SYSTEM; OPTIC TECTUM
CC (LAYERS D AND E OF SGFS), AMACRINE CELLS OF RETINA, NEURITES OF
CC DORSAL ROOT GANGLIA. ALSO EXPRESSED IN NONNEURONAL CELLS, E.G.
CC BLOOD VESSELS IN THE ENTIRE EMBRYO.
CC -!- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 NAM DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D45416; BAA09256.1; -.
InterPro: IPR000421; -.
InterPro: IPR000859; -.
InterPro: IPR000998; -.
Pfam: PF00431; CUB; 2.
Pfam: PF00754; F5_F8_typeC; 2.
Pfam: PF00629; NAM; 1.
PRINTS; PS0020; MAMDOMAIN.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00740; NAM_1; 1.
DR PROSITE; PS0060; NAM_2; 1.
KW Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
Cell adhesion.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 914 NEUROFILIN-1.
FT DOMAIN 20 847 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 848 870 POTENTIAL.
FT DOMAIN 871 914 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 139 CUB 1.
FT DOMAIN 145 263 CUB 2.
FT DOMAIN 273 422 F5/8 TYPE C 1.
FT DOMAIN 429 581 F5/8 TYPE C 2.
FT DOMAIN 636 801 MAM.
FT DISULFID 25 52 PROBABLE.
FT DISULFID 80 102 PROBABLE.
FT DISULFID 145 171 PROBABLE.
FT DISULFID 204 226 PROBABLE.
FT DISULFID 273 422 BY SIMILARITY.
FT DISULFID 429 581 BY SIMILARITY.
SQ SEQUENCE 914 AA; 102480 MW; DD2EE6D6F0CB68C CRC64;

Query Match 8.4%; Score 157; DB 1; Length 914;
Best Local Similarity 29.4%; Pred. No. 1.2e-05;
Matches 52; Conservative 27; Mismatches 68; Indels 30; Gaps 9;

```
Qy 50 IVTVSTNGSHSPRPHYPRNTVLVRLVAEENWVLIQTFDERFGLGLEDPEDDICKYDFV 109
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 29 IKILSPGLTSPGYSPQSPYHPKCEWLIQAPPEPYQIRIMFNPHFDLED RD---CKYDYV 85
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 110 EV--EPPSDGTILGRWCGSGTVPGKQISGNQIRIRFVSDYFPPSPGFCIHNYI--P 165
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 86 EVIDGDNAGRLWCKYCGK-IAPPLVSSGPFLEIKFVSD-YETHGAGFSIRVEVFRKP 143
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 166 Q-----FTEAVSPSLP--PSALPLDL-----LNNATAFTSLDLIRYLEPE 206
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 ECSRNETSSSGMIKSPGPEKYPNSLECTYIFAPKMSIILEFSEF-----LEPD 195
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 13
PCOL_HUMAN
ID PCOL_HUMAN STANDARD; PRT; 449 AA.
AC Q15113; O14550;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROCOLLAGEN C-PROTEINASE ENHANCER PROTEIN PRECURSOR (PCPE) (TYPE I
DE PROCOLLAGEN COOH-TERMINAL PROTEINASE ENHANCER) (TYPE 1 PROCOLLAGEN C-
DE PROTEINASE ENHANCER PROTEIN).
GN PCOLCE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Boutler J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirski R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: REQUIRED FOR NORMAL DORSAL DEVELOPMENT. TLD MAY INTERACT
 CC PHYSICALLY WITH DPP-C PROTEIN.
 CC -1- MISCELLANEOUS: MUTATIONS IN TLD GENE LEAD TO A PARTIAL
 CC TRANSFORMATION OF DORSAL ECTODERM INTO VENTRAL ECTODERM.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A (ZINC
 CC METALLOPROTEASE); ALSO KNOWN AS THE ASTACIN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 5 CUB DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M76976; AAA28491.1; -;
 DR EMBL; U04239; AAC46482.1; -;
 DR EMBL; AE003749; AAF56329.1; -;
 DR PIR; A39288; A39288.
 DR HSP; P00742; IFAX.
 DR MEROPS; M12.010; -;
 DR FlyBase; FBgn0003719; tld.
 DR InterPro; IPR000130; -;
 DR InterPro; IPR000152; -;
 DR InterPro; IPR000561; -;
 DR InterPro; IPR000859; -;
 DR InterPro; IPR001506; -;
 DR InterPro; IPR001881; -;
 DR Pfam; PF01400; Astacin; 1.
 DR Pfam; PF00431; CUB; 5.
 DR Pfam; PF00008; EGF; 2.
 DR PRINTS; PR00480; ASTACIN.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR PROSITE; PS01180; CUB; 5.
 DR PROSITE; PS00010; ASX_HYDROXYL; 2.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01187; EGF_CA; 2.
 KW Developmental protein; Hydrolase; Metalloprotease; Zinc; Glycoprotein;

KW EGF-like domain; Calcium; Signal; Repeat; Zymogen.
 FT SIGNAL 1 27 POTENTIAL.
 FT PROPEP 28 126
 FT CHAIN 127 1057 DORSAL-VENTRAL PATTERNING TOLLOID
 FT PROTEIN METALLOPROTEASE.
 FT DOMAIN 127 329
 FT DOMAIN 330 467
 FT DOMAIN 468 580
 FT DOMAIN 581 621 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 624 742 CUB.
 FT DOMAIN 743 783 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 787 899 CUB.
 FT DOMAIN 900 1016
 FT METAL 221 221 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 222 222 BY SIMILARITY.
 FT METAL 225 225 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 231 231 ZINC (CATALYTIC) (BY SIMILARITY).
 FT SITE 237 237 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 315 317 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 330 380 BY SIMILARITY.
 FT DISULFID 407 429 BY SIMILARITY.
 FT DISULFID 468 495 BY SIMILARITY.
 FT DISULFID 522 544 BY SIMILARITY.
 FT DISULFID 585 596 BY SIMILARITY.
 FT DISULFID 592 605 BY SIMILARITY.
 FT DISULFID 607 620 BY SIMILARITY.
 FT DISULFID 624 652 BY SIMILARITY.
 FT DISULFID 683 706 BY SIMILARITY.
 FT DISULFID 747 758 BY SIMILARITY.
 FT DISULFID 754 767 BY SIMILARITY.
 FT DISULFID 769 782 BY SIMILARITY.
 FT DISULFID 840 862 BY SIMILARITY.
 FT DISULFID 900 930 BY SIMILARITY.
 FT DISULFID 957 979 BY SIMILARITY.
 FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 431 431 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 533 533 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 634 634 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 667 667 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 781 781 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 854 854 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 908 908 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1057 AA; 120575 MW; 76F4B5AEB7996FBA CRC64;
 Query Match 8.0%; Score 148.5; DB 1; Length 1057;
 Best Local Similarity 38.0%; Pred. No. 7.7e-05;
 Matches 38; Conservative 20; Mismatches 35; Indels 7; Gaps 5;
 QY 50 ITVSTNGSIHSRPHPTYPRTNTVLVRLVAVENWVQLTDFDRFGLEDPEDDICKYDFV 109
 DB 472 LKLTQDSIDSPNYPMDYMPDKCEVWRITA-PDNHQVALKF-QSFELE--KHDGCAYDFV 527
 QY 110 EVEEE--PSDGTILGRWCSGTVPKQISKGNQIRIRFVSD 147
 DB 528 EIRDGNHSDSRLIGRCGDKLPNPK-TRSNQMYIRFVSD 566
 RESULT 15
 MAS2_HUMAN
 ID MAS2_HUMAN STANDARD; PRT; 686 AA.
 AC 000187; 075754;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE MANNAN-BINDING LECTIN SERINE PROTEASE 2 PRECURSOR (EC 3.4.21.-).
 DE (MANNOSE-BINDING PROTEIN ASSOCIATED SERINE PROTEASE 2) (MASP-2).
 GN MASP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

```

RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RX MEDLINE=9724212; PubMed=9087411;
RA Thiel S., Jensen T.V., Stover C.M., Schwaible W.J., Laursen S.B.,
RA Poulsen K., Willis A.C., Eggleton P., Hansen S., Holmskov U.,
RA Reid K.B.M., Jensenius J.C.;
RT "A second serine protease associated with mannan-binding lectin that
RL activates complement.";
RL Nature 386:506-510(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Thiel S., Vorup-Jensen T., Stover C.M., Schwaible W., Laursen S.B.,
RA Poulsen K., Willis A.C., Eggleton P., Hansen S., Holmskov U.,
RA Reid K.B.M., Jensenius J.C.;
RT "Identification and characterization of a novel protein of the human
RT complement system, mannan-binding lectin-associated serine protease-2
RT (MASP-2).";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: TRYPSIN PROTEASE THAT PRESUMABLY PLAYS AN IMPORTANT ROLE
CC IN THE INITIATION OF THE MBL COMPLEMENT ACTIVATION PATHWAY. AFTER
CC ACTIVATION IT CLEAVES C4 GENERATING C4A AND C4B.
CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Y09926; CAA71059.1; -
DR EMBL; X98400; CAA67050.1; -
DR HSP; P00763; LDPO.
DR MEROPS; S01.229; -.
DR InterPro; IPR000436; -
DR InterPro; IPR000859; -
DR InterPro; IPR001254; -
DR InterPro; IPR001314; -
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00084; sush1; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR Hydrolase; Complement pathway; Serine protease; Protease;
KW Glycoprotein; Sush1; Repeat; Signal; EGF-like domain; Hydroxylation.
FT SIGNAL 1 15
FT CHAIN 16 686
FT CHAIN 16 444
FT CHAIN 445 686
FT CHAIN 16 137
FT CHAIN 138 181
FT DOMAIN 184 296
FT DOMAIN 299 362
FT DOMAIN 365 431
FT ACT_SITE 483 483
FT ACT_SITE 532 532
FT ACT_SITE 633 633
FT MOD_RES 158 158
FT DISULFID 72 90
FT DISULFID 142 156
FT CUB-LIKE, CALCIUM-BINDING (POTENTIAL).
FT SUSHI 1.
FT SUSHI 2.
FT CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CHARGE RELAY SYSTEM (BY SIMILARITY).
FT HYDROXYLATION (POTENTIAL).
FT POTENTIAL.

```

```

FT DISULFID 152 165 POTENTIAL.
FT DISULFID 167 180 POTENTIAL.
FT DISULFID 184 211 POTENTIAL.
FT DISULFID 241 259 POTENTIAL.
FT DISULFID 300 348 POTENTIAL.
FT DISULFID 328 361 POTENTIAL.
FT DISULFID 366 412 POTENTIAL.
FT DISULFID 396 430 POTENTIAL.
FT DISULFID 434 552 INTERCHAIN (POTENTIAL).
FT DISULFID 598 618 POTENTIAL.
FT DISULFID 629 660 POTENTIAL.
SQ SEQUENCE 686 AA; 75685 MW; 4E34DED159448A2A CRC64;

Query Match 7.7%; Score 143.5; DB 1; Length 686;
Best Local Similarity 32.1%; Pred. No. 0.00012;
Matches 34; Conservative 25; Mismatches 42; Indels 5; Gaps 4;

QY 55 NGSIHSPRPHTYPRNTVLVWRLVAVENWVIOQLTFDEFGLEDDEDDICKYDFVEVEEP 114
DB 193 SGLSSPEYPRPYPKLSSCTYS-ISLEEGSVILDFVESFDVETHPTLCPYDLKIQ-- 249
QY 115 SDGTILGRWCGSGTVPKGQISKGNQIRFVSDYFPEPFCFHY 160
DB 250 TDREHGPFCGK-TLPHRIETKSNVTITFTWDE-SGDHTGWKIHY 293

```

Search completed: June 29, 2001, 17:11:59
Job time: 191 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2001, 17:07:48 ; Search time 27.91 Seconds
(without alignments)
1635.444 Million cell updates/sec

Title: US-09-457-066-2
Perfect score: 1858
Sequence: 1 MSLEGLLLTALAGROGT.....DVALEHHEECDCVCRGSTGG 345

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_16:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|--------------------|
| 1 | 1858 | 100.0 | 345 | 4 | Q9UL22 | Q9ul22 homo sapien |
| 2 | 1851 | 99.6 | 345 | 4 | Q9NRA1 | Q9nra1 homo sapien |
| 3 | 1667 | 89.7 | 345 | 11 | Q9QY71 | Q9qy71 mus musculu |
| 4 | 1649 | 88.8 | 345 | 11 | Q9BQX6 | Q9eqx6 rattus norv |
| 5 | 1638 | 88.2 | 345 | 11 | Q9JHV8 | Q9jvh8 mus musculu |
| 6 | 1617 | 87.0 | 345 | 13 | Q9I946 | Q9i946 gallus gall |
| 7 | 742.5 | 40.0 | 370 | 4 | Q9GZP0 | Q9gzp0 homo sapien |
| 8 | 736 | 39.6 | 370 | 11 | Q9QRT1 | Q9eqrl rattus norv |
| 9 | 183.5 | 9.9 | 3623 | 4 | O60494 | O60494 homo sapien |
| 10 | 182 | 9.8 | 691 | 13 | O57658 | O57658 gallus gall |
| 11 | 176 | 9.5 | 977 | 13 | Q9I925 | Q9i925 xenopus lae |
| 12 | 172 | 9.3 | 1012 | 11 | Q9WVM6 | Q9wvm6 mus musculu |
| 13 | 171 | 9.2 | 735 | 13 | O57381 | O57381 xenopus lae |
| 14 | 171 | 9.2 | 926 | 4 | Q9UQ00 | Q9uq00 homo sapien |
| 15 | 171 | 9.2 | 1015 | 4 | Q9Y6L7 | Q9y6l7 homo sapien |
| 16 | 171 | 9.2 | 3623 | 11 | O70244 | O70244 rattus norv |
| 17 | 169 | 9.1 | 241 | 11 | Q9Z135 | Q9z135 rattus norv |
| 18 | 164 | 8.8 | 3620 | 6 | Q9TU53 | Q9tu53 canis fami |
| 19 | 163.5 | 8.8 | 555 | 4 | Q9H2E2 | Q9h2e2 homo sapien |

| | | | | | | |
|----|-------|-----|------|----|--------|---------------------|
| 20 | 163.5 | 8.8 | 901 | 4 | Q9H2E4 | Q9h2e4 homo sapien |
| 21 | 163.5 | 8.8 | 901 | 4 | Q9H2D5 | Q9h2d5 homo sapien |
| 22 | 163.5 | 8.8 | 906 | 4 | Q9H2E3 | Q9h2e3 homo sapien |
| 23 | 163.5 | 8.8 | 906 | 4 | Q9H2D4 | Q9h2d4 homo sapien |
| 24 | 163.5 | 8.8 | 921 | 11 | Q9QX38 | Q9qxx38 rattus norv |
| 25 | 162.5 | 8.7 | 326 | 11 | O35251 | O35251 rattus norv |
| 26 | 159 | 8.6 | 1022 | 13 | O57460 | O57460 brachydanio |
| 27 | 158.5 | 8.5 | 704 | 4 | Q9H2E1 | Q9h2e1 homo sapien |
| 28 | 156 | 8.4 | 1008 | 13 | Q9DER7 | Q9der7 gallus gall |
| 29 | 153 | 8.2 | 415 | 4 | Q9UKZ9 | Q9ukz9 homo sapien |
| 30 | 153 | 8.2 | 1019 | 13 | O57382 | O57382 xenopus lae |
| 31 | 149 | 8.0 | 1013 | 4 | O43897 | O43897 homo sapien |
| 32 | 149 | 8.0 | 1013 | 4 | Q9NQS4 | Q9nqs4 homo sapien |
| 33 | 148 | 8.0 | 358 | 13 | O57434 | O57434 fugu rubrip |
| 34 | 148 | 8.0 | 423 | 13 | O57433 | O57433 fugu rubrip |
| 35 | 148 | 8.0 | 1013 | 11 | O62381 | O62381 mus musculu |
| 36 | 147 | 7.9 | 354 | 4 | O43915 | O43915 homo sapien |
| 37 | 146 | 7.9 | 746 | 5 | O01654 | O01654 halocynthia |
| 38 | 143.5 | 7.7 | 686 | 4 | Q9Y270 | Q9y270 homo sapien |
| 39 | 142.5 | 7.7 | 212 | 11 | O09020 | O09020 rattus norv |
| 40 | 142.5 | 7.7 | 701 | 11 | O9JJS9 | O9jjs9 rattus norv |
| 41 | 140 | 7.5 | 358 | 11 | P97946 | P97946 mus musculu |
| 42 | 139.5 | 7.5 | 1070 | 5 | P91972 | P91972 aplysia cal |
| 43 | 138.5 | 7.5 | 597 | 5 | Q26051 | Q26051 paracentrot |
| 44 | 137.5 | 7.4 | 699 | 4 | O95570 | O95570 homo sapien |
| 45 | 137.5 | 7.4 | 699 | 4 | Q9UF09 | Q9uf09 homo sapien |

ALIGNMENTS

RESULT 1
Q9UL22
ID Q9UL22 PRELIMINARY; PRT; 345 AA.
AC Q9UL22;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE SECRETORY GROWTH FACTOR-LIKE PROTEIN FALLOTEIN (SPINAL CORD-DERIVED GROWTH FACTOR).
DE HSCDGF.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RA Tsai Y.J., Lee R.K.K., Lin S.P.;
RT "Fallotein, a novel growth factor like gene identified in human uterus.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA MEDLINE=20317014; PubMed=10858496;
RA Hamada T., Oi-Tel K., Miyata Y.;
RT "A novel gene derived from developing spinal cords, SCDGF, is a unique member of the PDGF/VEGF family.";
RL FEBS Lett. 475:97-102(2000).
DR EMBL; AF091434; AAF00049.1; -;
DR EMBL; AB033831; BAB03286.1; -;
DR InterPro; IPR000072; -;
DR InterPro; IPR000859; -;
DR Pfam; PF00341; PDGF; 1.
DR Pfam; PF00431; PDGF; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS02078; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 39029 MW; CDE9E51F40633E78 CRC64;

Query Match 100.0%; Score 1858; DB 4; Length 345;

```

Best Local Similarity 100.0%; Pred. No. 2e-165;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFGLLLTSALAGORCTQAESNLSSKQFSSNKEQNGVODPOHERIITVSTNGSIHS 60
Db 1 MSFGLLLTSALAGORCTQAESNLSSKQFSSNKEQNGVODPOHERIITVSTNGSIHS 60

QY 61 PRFPHYPRNTVLVWRLVAVENWVQLTDFDERFGLDEDDICKYDFVEEPEPSDGTIL 120
Db 61 PRFPHYPRNTVLVWRLVAVENWVQLTDFDERFGLDEDDICKYDFVEEPEPSDGTIL 120

QY 121 GRWCGSTVPGKQISKGNOIRIRFVSDEYFPSEPGFCIHYNIYMPQFTEAVSPVLPPSA 180
Db 121 GRWCGSTVPGKQISKGNOIRIRFVSDEYFPSEPGFCIHYNIYMPQFTEAVSPVLPPSA 180

QY 181 LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWOLGKAFVFGKSRVVDLNL 240
Db 181 LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWOLGKAFVFGKSRVVDLNL 240

QY 241 LTEEVRLYSCTPRNFSVSIREELKRTDITFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
Db 241 LTEEVRLYSCTPRNFSVSIREELKRTDITFWPGCLLVKRCGNCACCLHNCNECQVPSK 300

QY 301 VTKYHEVLQRLPKTGVRLGKSLTDVALEHHEECDCVCRGSTG 345
Db 301 VTKYHEVLQRLPKTGVRLGKSLTDVALEHHEECDCVCRGSTG 345

RESULT 2
Q9NRA1 PRELIMINARY; PRT; 345 AA.
ID AC Q9NRA1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Li X., Ponten A., Aase K., Karlsson L., Abramsson A., Uutela M.,
RA Backstrom G., Hellstrom M., Bostrom H., Li H., Soriano P.,
RA Betsholtz C., Heldin C.-H., Alitalo K., Ostman A., Eriksson U.;
RT "PDGF-C is a novel protease-activated ligand for the PDGF alpha
RT receptor.";
RL Nat. Cell Biol. 0:0-0(2000).
DR EMBL; AF244813; AAF80597.1; -
DR InterPro; IPR000072; -
DR InterPro; IPR000859; -
DR Pfam; PF00341; PDGF.1.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 35043 MW; 590889CEA55C55EA CRC64;

Query Match 99.6%; Score 1851; DB 4; Length 345;
Best Local Similarity 99.4%; Pred. No. 9.1e-165;
Matches 343; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFGLLLTSALAGORCTQAESNLSSKQFSSNKEQNGVODPOHERIITVSTNGSIHS 60
Db 1 MSFGLLLTSALAGORCTQAESNLSSKQFSSNKEQNGVODPOHERIITVSTNGSIHS 60

QY 61 PRFPHYPRNTVLVWRLVAVENWVQLTDFDERFGLDEDDICKYDFVEEPEPSDGTIL 120
Db 61 PRFPHYPRNTVLVWRLVAVENWVQLTDFDERFGLDEDDICKYDFVEEPEPSDGTIL 120

QY 121 GRWCGSTVPGKQISKGNOIRIRFVSDEYFPSEPGFCIHYNIYMPQFTEAVSPVLPPSA 180
Db 121 GRWCGSTVPGKQISKGNOIRIRFVSDEYFPSEPGFCIHYNIYMPQFTEAVSPVLPPSA 180

```

```

Db 121 GRWCGSTVPGKQISKGNOIRIRFVSDEYFPSEPGFCIHYNIYMPQFTEAVSPVLPPSA 180
QY 181 LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWOLGKAFVFGKSRVVDLNL 240
Db 181 LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWOLGKAFVFGKSRVVDLNL 240
QY 241 LTEEVRLYSCTPRNFSVSIREELKRTDITFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
Db 241 LTEEVRLYSCTPRNFSVSIREELKRTDITFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
QY 301 VTKYHEVLQRLPKTGVRLGKSLTDVALEHHEECDCVCRGSTG 345
Db 301 VTKYHEVLQRLPKTGVRLGKSLTDVALEHHEECDCVCRGSTG 345

RESULT 3
Q9QY71 PRELIMINARY; PRT; 345 AA.
ID AC Q9QY71;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-WAR-2001 (Tremblrel. 16, Last annotation update)
DE FALLOTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RA Tsai Y.-J., Lee R.K.-K., Chen Y.-H., Lin S.-P., Cheng W.T.-K.;
RT "cDNA cloning of follotain from mouse ovary.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF117608; AAF22516.1; -
DR InterPro; IPR000072; -
DR InterPro; IPR000859; -
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38741 MW; 3A58A1F701B84EA2 CRC64;

Query Match 89.7%; Score 1667; DB 11; Length 345;
Best Local Similarity 87.0%; Pred. No. 1.4e-147;
Matches 300; Conservative 27; Mismatches 18; Indels 0; Gaps 0;

QY 1 MSFGLLLTSALAGORCTQAESNLSSKQFSSNKEQNGVODPOHERIITVSTNGSIHS 60
Db 1 MSFGLLLTSALAGORCTQAESNLSSKQFSSNKEQNGVODPOHERIITVSTNGSIHS 60

QY 61 PRFPHYPRNTVLVWRLVAVENWVQLTDFDERFGLDEDDICKYDFVEEPEPSDGTIL 120
Db 61 PRFPHYPRNTVLVWRLVAVENWVQLTDFDERFGLDEDDICKYDFVEEPEPSDGTIL 120

QY 121 GRWCGSTVPGKQISKGNOIRIRFVSDEYFPSEPGFCIHYNIYMPQFTEAVSPVLPPSA 180
Db 121 GRWCGSTVPGKQISKGNOIRIRFVSDEYFPSEPGFCIHYNIYMPQFTEAVSPVLPPSA 180

QY 181 LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWOLGKAFVFGKSRVVDLNL 240
Db 181 LPDLNNAITAFSTLEDLIRYLEPERWQDLEDLYRPTWOLGKAFVFGKSRVVDLNL 240

QY 241 LTEEVRLYSCTPRNFSVSIREELKRTDITFWPGCLLVKRCGNCACCLHNCNECQVPSK 300
Db 241 LTEEVRLYSCTPRNFSVSIREELKRTDITFWPGCLLVKRCGNCACCLHNCNECQVPSK 300

QY 301 VTKYHEVLQRLPKTGVRLGKSLTDVALEHHEECDCVCRGSTG 345
Db 301 VTKYHEVLQRLPKTGVRLGKSLTDVALEHHEECDCVCRGSTG 345

```

```
RESULT 4
Q9EQX6 PRELIMINARY; PRT; 345 AA.
AC Q9EQX6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE SPINAL CORD-DERIVED GROWTH FACTOR.
GN RSCDGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR: TISSUE=KIDNEY;
RA Hamada T., Ui-Tei K., Imaki J., Miyata Y.;
RT "Molecular Cloning of SCDF-B, a Novel Growth Factor Homologous to
RL SCDF/PDGF-C/fallotin.";
RL Biochem. Biophys. Res. Commun. 0:0-0(2000).
DR EMBL; AB033830; BAB19969.1; -.
SQ SEQUENCE 345 AA; 38734 MW; F296DA6E9B765D10 CRC64;

Query Match 88.8%; Score 1649; DB 11; Length 345;
Best Local Similarity 85.8%; Pred. No. 6.8e-146;
Matches 296; Conservative 29; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSLFGLLLTSALAGROGTOAESNLSSKQFSSNKQNGVQDPQHERITVTNGSIHS 60
DB 1 MLLGLLLTSALAGROGTGRAESNLSSKLQSSDKQNGVQDPQHERVVTISNGSIHS 60
QY 61 PRPHTYPRNTVLVRLVAVENWVQLTFDERFGLDEPDDICKYDFVEVEEPSDGTIL 120
DB 61 PRPHTYPRNTVLVRLVAVENWVQLTFDERFGLDEPDDICKYDFVEVEEPSDGSVL 120
QY 121 GRWCGSTVPKGQISGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQTEAVSPSVLPSSA 180
DB 121 GRWCGSTVPKGQISGNQIRIRFVSDEYFPSEPGFCIHYSIMPQVTTETSPSVLPSS 180
QY 181 LPDLNNATFSTLEDLIRYLEPERWQLDLEDLYRPTWOLLGKAFVGRKSRVVDLNL 240
DB 181 LSLLNNAVTAFSTLEELIRYLEPDRWQDLDLSLYKPTWLLGKAFLYVKKSKAVNLNL 240
QY 241 LEEVRLYSCPTPNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQCVPSK 300
DB 241 LKEEVKLYSCPTPNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQCVPRK 300
QY 301 VTKKYHEVLQRPKTVGRGLHKLSTDLVALEHHECDVCVRGSGTG 345
DB 301 VTKKYHEVLQRPKTVGRGLHKLSTDLVALEHHECDVCVRGNAGG 345

RESULT 5
Q9JHV8 PRELIMINARY; PRT; 345 AA.
AC Q9JHV8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PLATELET-DERIVED GROWTH FACTOR C.
GN PDGFC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS-WEBSTER/NIH;
RA Ding H., Wu X., Kim I., Tam P.P.L., Koh G.Y., Nagy A.;
RT "The mouse pdgfc gene: Dynamic expression in embryonic tissues during
RT organogenesis.";
RL Mech. Dev. 0:0-0(2000).
```

```
DR EMBL; AF286725; AAF91483.1; -.
DR InterPro; IPR000072; -.
DR InterPro; IPR000859; -.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS0278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38886 MW; FA1486BED6D362F8 CRC64;

Query Match 88.2%; Score 1638; DB 11; Length 345;
Best Local Similarity 85.8%; Pred. No. 7.2e-145;
Matches 296; Conservative 27; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSLFGLLLTSALAGROGTOAESNLSSKQFSSNKQNGVQDPQHERITVTNGSIHS 60
DB 1 MLLGLLLTSALAGROGTGRAESNLSSKLQSSDKQNGVQDPQHERVVTISNGSIHS 60
QY 61 PRPHTYPRNTVLVRLVAVENWVQLTFDERFGLDEPDDICKYDFVEVEEPSDGTIL 120
DB 61 PKFHTYPRNMVLRVAVDENVRLTQTFDERFGLDEPDDICKYDFVEVEEPSDGSVL 120
QY 121 GRWCGSTVPKGQISGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQTEAVSPSVLPSSA 180
DB 121 GRWCGSTVPKGQISGNQIRIRFVSDEYFPSEPGFCIHYSIMPQVTTETSPSVLPSS 180
QY 181 LPDLNNATFSTLEDLIRYLEPERWQLDLEDLYRPTWOLLGKAFVGRKSRVVDLNL 240
DB 181 LSLLNNAVTAFSTLEELIRYLEPDRWQDLDLSLYKPTWLLGKAFLYVKKSKVNLNL 240
QY 241 LEEVRLYSCPTPNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQCVPSK 300
DB 241 LKEEVKLYSCPTPNFSVSIREELKRTDTIFWPGCLLVKRCGNCACCLHNCNECQCVPRK 300
QY 301 VTKKYHEVLQRPKTVGRGLHKLSTDLVALEHHECDVCVRGSGTG 345
DB 301 VTKKYHEVLQRPKTVGRGLHKLSTDLVALEHHECDVCVRGNAGG 345

RESULT 6
Q91946 PRELIMINARY; PRT; 345 AA.
AC Q91946;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE SPINAL CORD-DERIVED GROWTH FACTOR.
GN SCDFG.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=SPINAL CORD;
RA MEDLINE-20317014; PubMed-10858496;
RA Hamada T., Ui-Tei K., Miyata Y.;
RT "A novel gene derived from developing spinal cords, SCDFG, is a unique
RT member of the PDGF/VEGF family.";
RL FEBS Lett. 475:97-102(2000).
DR EMBL; AB033829; BAB03265.1; -.
DR InterPro; IPR000072; -.
DR InterPro; IPR000859; -.
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS0278; PDGF_2; 1.
DR SMART; SM00042; CUB; 1.
SQ SEQUENCE 345 AA; 38940 MW; 97ACEA992BF5128C CRC64;

Query Match 87.0%; Score 1617; DB 13; Length 345;
Best Local Similarity 84.1%; Pred. No. 6.6e-143;
```



```
Query Match          9.2%; Score 171; DB 13; Length 735;
Best Local Similarity 43.3%; Pred. No. 2.2e-07;
Matches 45; Conservative 14; Mismatches 37; Indels 8; Gaps 5;

QY 55 NGSIHSPRPHTYPRNTVLVRLVAEENWVWVQVVAQAQ-YRISLQF-EVELEG--NDVCKYDFVEVRSG 113
DB 589 NGSITSPGWPKEYPTNKNVQVVAQAQ-YRISLQF-EVELEG--NDVCKYDFVEVRSG 644

QY 114 -PSDGTILGRWCSTGTPVKQISKGNQIRIRFVSDYFPSPGFCIH 156
DB 645 LPSDKLHGRFCGSET-ELPAVITSQYNNMRVFEKSDNTV-SKRGF 686

RESULT 14
ID Q9UQ00 PRELIMINARY; PRT; 926 AA.
AC Q9UQ00;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE KIAA0932 PROTEIN (FRAGMENT).
GN KIAA0932.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirose M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:63-70(1999).
DR EMBL; AB023149; BAA76776.1; -.
DR HSSP; P00736; IAPQ.
DR InterPro; IPR000130; -.
DR InterPro; IPR000152; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR000859; -.
DR InterPro; IPR001506; -.
DR InterPro; IPR001881; -.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF01400; Astacin; 1.
DR PRINTS; PR00480; ASTACIN.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SMART; SM00042; CUB; 1.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
FT NON_TER
SQ SEQUENCE 926 AA; 103873 MW; 58FA35CDDDE10970B CRC64;

Query Match          9.2%; Score 171; DB 4; Length 926;
Best Local Similarity 43.5%; Pred. No. 3e-07;
Matches 47; Conservative 14; Mismatches 39; Indels 8; Gaps 6;

QY 55 NGSIHSPRPHTYPRNTVLVRLVAEENWVWVQVVAQAQ-YRISLQF-EVELEG--NDVCKYDFVEVRSG 113
DB 539 NGTITSPGWPKEYPTNKNVQVVAQAQ-YRISLQF-EVELEG--NDVCKYDFVEVRSG 594

QY 114 -PSDGTILGRWCSTGTPVKQISKGNQIRIRFVSDYFPSPGFCIH 160
DB 595 LSPDAKLHGRFCGSET-PEVITSQNNMRVFEKSDNTV-SKRGFRAHF 640
```

```
RESULT 15
Q9Y6L7 PRELIMINARY; PRT; 1015 AA.
AC Q9Y6L7;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE TOLLOID-LIKE 2 PROTEIN.
GN TLL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Scott I.C., Greenspan D.S.;
RT "Sequence of the human mammalian tolloid-like 2 (mTll-2) and
RT chromosomal localisation of the cognate gene TLL2.";
RL Dev. Biol. 0:0-0(1999).
DR EMBL; AF059516; AAD42979.1; -.
DR HSSP; P00736; IAPQ.
DR InterPro; IPR000130; -.
DR InterPro; IPR000152; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR000859; -.
DR InterPro; IPR001506; -.
DR InterPro; IPR001881; -.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF01400; Astacin; 1.
DR PRINTS; PR00480; ASTACIN.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SMART; SM00042; CUB; 1.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 1015 AA; 113536 MW; 25F5B23065861593 CRC64;
```

Query Match 9.2%; Score 171; DB 4; Length 1015;
Best Local Similarity 43.5%; Pred. No. 3.4e-07;
Matches 47; Conservative 14; Mismatches 39; Indels 8; Gaps 6;

```
QY 55 NGSIHSPRPHTYPRNTVLVRLVAEENWVWVQVVAQAQ-YRISLQF-EVELEG--NDVCKYDFVEVRSG 113
DB 628 NGTITSPGWPKEYPTNKNVQVVAQAQ-YRISLQF-EVELEG--NDVCKYDFVEVRSG 683

QY 114 -PSDGTILGRWCSTGTPVKQISKGNQIRIRFVSDYFPSPGFCIH 160
DB 684 LSPDAKLHGRFCGSET-PEVITSQNNMRVFEKSDNTV-SKRGFRAHF 729
```

Search completed: June 29, 2001, 17:11:34
Job time: 226 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2001, 00:39:28 ; Search time 2381.26 Seconds
(without alignments)
11458.266 Million cell updates/sec

Title: US-09-457-066-1
Perfect score: 1764
Sequence: 1 attatgtgaaactaccctg.....cttgtgtgtgtgatagga 1764

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues 2688314
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_ba3.*
4: gb_in1.*
5: gb_in2.*
6: gb_in3.*
7: gb_om.*
8: gb_ov.*
9: gb_pat1.*
10: gb_pat2.*
11: gb_ph.*
12: gb_pl1.*
13: gb_pl2.*
14: gb_pl3.*
15: gb_pl4.*
16: gb_ba1.*
17: em_ba2.*
18: em_fun.*
19: em_htgo_hum.*
20: em_htgo_inv.*
21: em_htgo_rod.*
22: em_htg_hum1.*
23: em_htg_hum2.*
24: em_htg_hum3.*
25: em_htg_hum4.*
26: em_htg_hum5.*
27: em_htg_hum6.*
28: em_htg_hum7.*
29: em_htg_hum8.*
30: em_htg_inv1.*
31: em_htg_inv2.*
32: em_htg_other.*
33: em_htg_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in.*
42: em_om.*
43: em_or.*

em_ov.*
44: em_pat.*
45: em_ph.*
46: em_pl.*
47: em_pl.*
48: em_ro.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vi.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vil.*
59: gb_vil.*
60: gb_vil.*
61: gb_vil.*
62: gb_vil.*
63: gb_vil.*
64: gb_vil.*
65: gb_vil.*
66: gb_vil.*
67: gb_vil.*
68: gb_vil.*
69: gb_vil.*
70: gb_vil.*
71: gb_vil.*
72: gb_vil.*
73: gb_vil.*
74: gb_vil.*
75: gb_vil.*
76: gb_vil.*
77: gb_vil.*
78: gb_vil.*
79: gb_vil.*
80: gb_vil.*
81: gb_vil.*
82: gb_vil.*
83: gb_vil.*
84: gb_vil.*
85: gb_vil.*
86: gb_vil.*
87: gb_vil.*
88: gb_vil.*
89: gb_vil.*
90: gb_vil.*
91: gb_vil.*
92: gb_vil.*
93: gb_vil.*
94: gb_vil.*
95: gb_vil.*
96: gb_vil.*
97: gb_vil.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Query Length | DB ID | Description |
|------------|--------|-------|--------------|-------|--------------------|
| 1 | 1746 | 99.0 | 1760 | 9 | AX044518 Sequence |
| 2 | 1734 | 98.3 | 3007 | 88 | AF091434 Homo sapi |
| 3 | 1732.4 | 98.2 | 2849 | 9 | AX047650 Sequence |
| 4 | 1701 | 96.4 | 2152 | 89 | AF244813 Homo sapi |
| 5 | 1574 | 89.2 | 1817 | 85 | AB033831 Homo sapi |
| 6 | 1035 | 58.7 | 1035 | 9 | AX027935 Sequence |
| 7 | 1035 | 58.7 | 1035 | 91 | AX028032 Sequence |
| 8 | 883.2 | 50.1 | 3571 | 9 | AX044520 Sequence |

```
9 881.6 50.0 2692 94 AF117608 Mus muscu
10 780 44.2 1116 94 AB033830 Rattus no
11 779.8 44.2 1675 8 AB033829 Gallus ga
12 773 43.8 1038 94 AF286725 Mus muscu
13 662.4 37.6 191314 61 AC009582 Homo sapi
14 447 25.3 504 9 AX027960 Sequence
15 447 25.3 504 91 AX028057 Sequence
16 317.8 18.0 82113 63 AC015451 Homo sapi
17 283.4 16.1 289 9 AX027970 Sequence
18 283.4 16.1 289 9 AX027993 Sequence
19 283.4 16.1 289 56 AX028067 Sequence
20 283.4 16.1 289 56 AX028090 Sequence
21 279 15.8 279 9 AX027968 Sequence
22 279 15.8 279 9 AX027989 Sequence
23 279 15.8 279 56 AX028065 Sequence
24 279 15.8 279 56 AX028086 Sequence
25 278.6 15.8 284 9 AX027963 Sequence
26 278.6 15.8 284 9 AX027982 Sequence
27 278.6 15.8 284 56 AX028060 Sequence
28 278.6 15.8 284 56 AX028079 Sequence
29 273 15.5 300 9 AX027962 Sequence
30 273 15.5 300 9 AX027981 Sequence
31 273 15.5 300 56 AX028059 Sequence
32 273 15.5 300 56 AX028078 Sequence
33 272.4 15.4 152779 64 AC015837 Homo sapi
34 261 14.8 261 9 AX027967 Sequence
35 261 14.8 261 9 AX027988 Sequence
36 261 14.8 261 56 AX028064 Sequence
37 261 14.8 261 56 AX028085 Sequence
38 259.4 14.7 275 9 AX027964 Sequence
39 259.4 14.7 275 9 AX027984 Sequence
40 259.4 14.7 275 56 AX028061 Sequence
41 259.4 14.7 275 56 AX028081 Sequence
42 255 14.5 278 9 AX027965 Sequence
43 255 14.5 278 9 AX027986 Sequence
44 255 14.5 278 56 AX028062 Sequence
45 255 14.5 278 56 AX028083 Sequence

ALIGNMENTS

RESULT 1
AX044518
LOCUS AX044518 1760 bp DNA
DEFINITION Sequence 32 from Patent WO00066736.
ACCESSION AX044518
VERSION AX044518.1 GI:11343373
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1760)
AUTHORS Gilbert,T., Hart,C.E., Sheppard,P.O. and Gilbertson,D.G.
TITLE Growth factor homolog zveg4
JOURNAL Patent: WO 0066736-A 32 09-NOV-2000.
ZymoGenetics, Inc. (US)
FEATURES
    Location/Qualifiers
        1..1760
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            154..1191
                /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAC11710.1"
            /db_xref="GI:11343374"
            /translation="MSLFGILLTTSALAGQRGQTQAESNLSSKFORSSNKQGVQDP
            QHERLITVNGSTHSRPFTYPRNTVLVRLVAEENVIQLTFDERFGLDEPDD
            ICKYDFVEVPSDGTILGRNCGSGTVPGKQISKQNIIRIRVSEYPSRFGCIHY
            NIVMPQTEAVSPVLPSPALPLDNLNNAITAFSTLEDLIRYLEPFWOLDLEDIRP
            TWQLGKAFVGRKSRVVDNLNLTVEEVLRY SCTPRNFVSVIREELKRTDITFWPGCLL
            VKRGGNCACCLHNCQCQVPSKVKRYKHYEVLQLRPRTGTGRGLHKLSTDLVALEHHEE
```

QY 961 ttctgcccagggttgctctctcgtgtaaacgctgtggtggtggaactgtgctgtgtgtctcacc 1020
Db 961 TTCTGGCCAGGTTGCTCTCTCGTTAAACGCTGTGTGGGAACTGTGCTGTGTCTCCAC 1020
QY 1021 aattgcaatgaatgcaatgtcccaagcaaaagttaactaaaaataaccagaggtcctt 1080
Db 1021 AATTGCAATGAATGTCAATGTGCTCCCAAGCAAAAGTTACTAAAAAATACCAGAGGTCTT 1080
QY 1081 cagttgagaccaaagaccggtgtcaggggattgcacaaaatacctcaccagcgtggccctg 1140
Db 1081 CAGTTGAGACCAAGACCGGTGTCCAGGGATTGCACAAATCACTCACCGAGTGGCCCTG 1140
QY 1141 gagcacatagagagtgactgtgtgagaggggagcagagagatagccgcacatc 1200
Db 1141 GAGCACATGAGAGTGTGACTGTGTGAGAGGGAGACAGAGAGATAGCCGATCAC 1200
QY 1201 caccagcagctctgccagagctgtcagtllyllycagtggtcgttctattagagaac 1260
Db 1201 CACCAGCAGCTCTGCCAGAGCTGTGCAGT---GCAGTGGCTGATTTCTATTAGAGAAC 1260
QY 1261 gtagcgttatctccatcccttaactcagttgtttgtctcaaggacctttcatcttcagg 1320
Db 1261 GTATGCGTTATCTCCATCTTAATCTCAGTTGTTGTCTCAAGGACCTTTCATCTTCAGG 1320
QY 1321 attacagtgcatctgaaagagagagacatacaacagaattagagtggtgcaacagctc 1380
Db 1321 AATTACAGTGCAATCTCAAAAGAGAGAGATCAAAACAGAAATTAGCAGTTGTGCAACAGCTC 1380
QY 1381 ttttgagagggcctaaagacagaggaagaaaggctctcaatcgtgaaagaaaaataaa 1440
Db 1381 TTTTGAGAGGGCCCTAAAGACAGGAGAAAGGCTTTCAATCTGTGAAAGAAAAATAAA 1440
QY 1441 tgtgtatataatagatcacacagctagtttcagagttaccatgtacatgtattccactagct 1500
Db 1441 TGTGTATTAATAATAGATCACACAGCTAGTTTTCAGAGTTACCATGTACCATTTTCCATGCT 1500
QY 1501 ggggtctgtatttcagttcttcgatacagcgttagggtaattgtcagtacagagaaataaac 1560
Db 1501 GGGTCTGTATTTCAGTTCTTCGATACGGCTTAGGGTAAATGTGTCAGTACAGGAAAAAAC 1560
QY 1561 tbtgcaagtgtgacacctgtatccgtgtcctgtcttaactctaaagctccatgtcctgggc 1620
Db 1561 TGTGCAAGTGAGCACCTGATTCGGTTGCCCTTGTCTTAACCTTAAAGCTCCCATGTCTCGGC 1620
QY 1621 ctataatcatataaatcgtgatt 1680
Db 1621 CTAAAACTGTATATAATCTGGATTTTTTTTTTTTTTTTTTTTTTTTTTTTGTCTCATATTCAT 1680
QY 1681 accagaaatctctgtactacaaacacctgggttttttaaaaggaactatgtgtctatgaat 1740
Db 1681 ACCAGAACTCTGTACTACAAACCTGGTTTTTAAAGGAACATGTGTGCTATGAAT 1740
QY 1741 taaactgtgtcgtgtgatagga 1764
Db 1741 TAAACTGTGTGCTGTGTATAGGA 1760

RESULT 2
AF091434 3007 bp mRNA PRI 22-JUN-2000
LOCUS Homo sapiens secretory growth factor-like protein fallotein mRNA, complete cds.
DEFINITION
ACCESSION AF091434
VERSION AF091434.1 GI:6002592
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3007)
AUTHORS Tsai,Y., Lee,R.K., Lin,S. and Chen,Y.
TITLE Identification of a novel platelet-derived growth factor-like gene, fallotein, in the human reproductive tract

JOURNAL Biochim. Biophys. Acta 1492 (1), 196-202 (2000)
MEDLINE 20461776
REFERENCE 2 (bases 1 to 3007)
AUTHORS Tsai,Y., Lee,R.K.K. and Lin,S.P.
TITLE Direct Submission
JOURNAL Submitted (14-SEP-1998) Dept. Medical Research, Mackay Memorial Hospital, 45 Min Sheng Road, Tamshui, Taipei County 25115, Taiwan
FEATURES
source Location/Qualifiers
1. 3007
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="uterus"
492..1529
/evidence=not_experimental
/product="secretory growth factor-like protein fallotein"
/protein_id="AAF00049.1"
/db_xref="GI:6002593"
/translation="MSLFGLLLLTSALAGRQGTQAESNLSKFKQSSNKQNGVQDP
QHERIITVSTNGSIHSPRPHTYPRNTVLVWLVAVEENVMVQLTFDFRFGLEPEDD
ICKYDFVEVEPSPDGTILGRWCGSGTVPKQISKGNOIRFVSDVFPSPGFCIH
NIVMPQTEAVSPSVLPDLSALPLDILNNAITAFSTLEDLIRYLPERQWLDLEDLYR
TWQLGKAFVGRKSRVVLDNLLETEVRLYISCTPNFVSIRELKRTDITFWPGCLL
VKRCGNCACCLHNCNECCQVPSKVKYKHYEVLQLRPKTYGRGLHLSLTDVALEHHEE
CDCVCRGSTGG"
polya_signal 858 a 613 c 635 g 901 t
BASE COUNT 2974..2979
ORIGIN
Query Match 98.3%; Score 1734; DB 88; Length 3007;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1759; Conservative 0; Mismatches 0; Indels 5; Gaps 2;
QY 1 attatgtgaaactaccctggattctctgtgcagagcaggtcggcgctccacccc 60
Db 339 ATTATGTGAAACTACCCTGCGATTCTCTGTGCCAGAGCAGGCTCGGGCTTCCACCCC 398
QY 61 agtgagccttccctggcgtggtgaaagagactcggagtcgctgtccaaagtgc 120
Db 399 AGTGAGCGCTTCCCTGGCGTGGTGAAGAGACTCGGGAGTCGCTGTCCAAAGTGCC 458
QY 121 cgcctgagtgagctctcaccagtcagccaaatgagcctcttcgggcttctcctgctg 180
Db 459 CGCGTGAGTGAGCTCTCACCACCTCAGCCAAATGAGCCTTTCGGGCTTCTCTGCTG 518
QY 181 acatgcctcggcggccagacagagggactcaggcggaatcccaacctgagtagtaaa 240
Db 519 ACATCTGCCCTGGCGGCCAGAGACAGGGAGCTCAGGGGGAATCCAACTGAGTAGTAA 578
QY 241 ttccagtttccagcaacaagaaacagagcagaggtacagatcctcagcatgagagaatt 300
Db 579 TTCCAGTTTCCAGCAACAAGAACAGAACCGAGGTACAGATCCTCAGCATGAGAGAATT 638
QY 301 attactgtctactaatggaagtattcacagcccaagtttccctcactattatccaaga 360
Db 639 ATTACTGTCTACTTAATGGAAGTATTACAGCCCAAGGTTTCCCTACTACTTATCCAGA 698
QY 361 aatacggcttggtatgagattagtagcagtagaggaataatgtaggatacaactacg 420
Db 699 AATAGGCTTGGTATGGAGATTAGTAGCAGTAGAGGAAATGTATGGATACACTTACG 758
QY 421 ttgtatgaagatttggcttgaagaccagaagatgacatatgaagtatgattttgta 480
Db 759 TTTGATGAAGATTGGGCTTGAAGACCCAGAAAGATGACATATGCAAGTATGATTTTCTA 818
QY 481 gaagttaggaaccagtgatggaaactatattagggcgtggtggttctggtactgta 540
Db 819 GAACTTAGGAACCCAGTATGGAACCTATATAGGGCCCTGGTGTGGTCTGGTACTGTA 878
QY 541 ccaggaacacagatttctcaaggaaatacaattaggaataagattgtctctgatgaat 600
Db 879 CCAGGAAACAGATTCTTAAAGGAATCAATATGAGTAAGATTGTATCTGTGTAATAT 938

QY 601 ttctcttgaacagggttctgcatccactacatacaacattgtcatgccacaatttcagaa 660
Db 939 TTTCCTCTGAACAGGGTCTGCATCCACATACAACATTGTGCATGCCACAATTCAGAA 998
QY 661 gctgtgagctcctcagtgctaccccccttcagcttggcaactggagcctgcttaaatgct 720
Db 999 GCTGTGAGTCCCTTCAGTGTCTACCCCTTCAGCTTTGGCCACTGGACCTGCTTAATGCT 1058
QY 721 ataactgcttagtacccttgagagaccttattcgatatcttgaaccagagagatggcag 780
Db 1059 ATAAGTGCCTTTAGTACCTTGGAGACCTTATTCGATATCTTGAACAGAGAGATGGCAG 1118
QY 781 ttggacttagaagatctatataggccaaacttggcaactcttctggcaaggcttttggtttt 840
Db 1119 TTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGCTTTTGTGTTT 1178
QY 841 ggaagaaatccagagtggtgagctggaaccttctaagagagagaggttaagattatcacgc 900
Db 1179 GGAAGAAATCCAGAGTGGTGTGATCTGAACCTTCTTAACAGAGAGAGGTAAAGATTATACGC 1238
QY 901 tgcacactgtaacttctcagtgctccataaggaagaaactaaagagaaacagcattaccatt 960
Db 1239 TGCACACTGTAACCTTCTCAGTGTCCATAGGGGAAGAACTTAAGAGAACCCGATACCAT 1298
QY 961 ttctggccaggttctctctgtttaaacgctgtgtgggaactgtgcctgttggctccac 1020
Db 1299 TTCGGCCAGGTGCTCTCTGTTAAACGCTGTGTGGGAACTGTGCTGTGTCTCCAC 1358
QY 1021 aattgcaatgaatgcaatgtgtcccaagcaaaagttaataaaaaataccacaggtcctt 1080
Db 1359 AATTGCAATGAATGCAATGTGTCCCAAGCAAAAGTTACTTAAAAATACACAGGTCCTT 1418
QY 1081 cagttgagacaaagacggtgtcagggattgcacaaatcactcaccagcgtgcccctg 1140
Db 1419 CAGTTGAGACCAAGACCGGTGTCCAGGGATTGCACAAATCACTCACCAGCTGGCCCTG 1478
QY 1141 gagcaccatgagggtgtgactgtgtgcagagggagcacagaggagatagccgcatcac 1200
Db 1479 GAGCACCATGAGGAGTGTACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCCGCATCAC 1538
QY 1201 caccagagctcttgcacagagctgtgcagtllygcagtggtgcttattatagagaac 1260
Db 1539 CACAGCAGCTCTTGCACAGAGCTGTGCAGT- ---GCAGTGGCTGATTCTATTAGAGAAC 1594
QY 1261 gtatgcttattccatcccttaactcagttgtttgttgcagagcctttcatctcagg 1320
Db 1595 GTATGCGTATTCTCCATCCCTTAATCTCAGTTGTGTTCAGGAGACCTTTCATCTTCAGG 1654
QY 1321 atttacagtgctattctgaaagagagacatcaaacagaaattaggagttgtgcaacagctc 1380
Db 1655 ATTTACAGTGCATTTCAAGAGAGAGACATCAACAGAAATTAGGAGTTGTGCAACAGCTC 1714
QY 1381 ttttgagagaggcctaaagagacaggagaaaggtcttcaactcgtggaaagaaataaa 1440
Db 1715 TTTTGAGAGAGGGCTTAAAGACAGGAGAAAGGTCTTCAATCGTGAAGAAATATAA 1774
QY 1441 tgttgtaataatagatcaccagctagtctcagagttaccatgtacgtattccactagct 1500
Db 1775 TGTGTGTAATAATAGATCACCAGTAGTTCAGAGTTTACCATGTACGTATTTCACATAGCT 1834
QY 1501 ggggtctgtatttcagttcttcgatcagcgttagggtaattgtcagtcacagaaaaaac 1560
Db 1835 GGGTCTGTATTTCAGTTCCTTCGATACGGGCTTAGGGTAATGTAGTACAGGAAAAAAC 1894
QY 1561 tgtgcaagtgagcacctgattccgttgccttgaacttaacttaagcctcctgctgggc 1620
Db 1895 TGTGCAAGTGAGCACCTGATTCCGCTTGCCTTGAACCTTAAAGCTCCATGCTCCTGGC 1954
QY 1621 ctaaaatcgataaaaatctggattttttttttttttttttttttttttttttttttttta 1680
Db 1955 CTAATAATCGTATAAAATCTGGA- TTTTTTTTTTTTTTTTTTTTTTTTCTCATATATGTA 2013

QY 1681 accagaaacttctatgtactacaaacctgtttttaaaaggaactatgttctatgaat 1740
Db 2014 ACCAGAACTTCTATGTACTACAACTGCTTTTTTAAAGGAACATATGTGTCTATGAAT 2073
QY 1741 taaaacttgctgctgctgataga 1764
Db 2074 TAAACTGTGCTGCTGCTGATAGA 2097
RESULT 3
AX047650 2849 bp DNA PAT 15-DEC-2000
LOCUS Sequence 9 from Patent WO0070050.
DEFINITION AX047650
ACCESSION AX047650
VERSION AX047650.1 GI:11876693
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2849)
AUTHORS Baker, K.P., Chen, J., Ferrara, N., Fong, S., Goddard, A., Gurney, A.L., Hillan, K.J., Kuo, S.S., Tumas, D., and Wood, W.I.
TITLE Compositions and methods for the treatment of immune related diseases
JOURNAL Patent: WO 0070050-A 9 23-NOV-2000;
FEATURES Location/Qualifiers
source 1..2849
/organism="Homo sapiens"
/db_xref="taxon:9606"
2715
unsure /note="unknown base"
BASE COUNT 851 a 528 c 619 g 850 t 1 others
ORIGIN

Query Match 98.2%; Score 1732.4; DB 9; Length 2849;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

QY 1 attatggaaactaccctgcgattctctgtccagagcaggctcggtccaccoccc 60
Db 132 ATTATGTGAAACTACCTCGCATTCCTGTCTGCCAGAGCAGGCTGGCGCTTCCACCCC 191
QY 61 agtgcagcttcccttgcgggtgtgtaagagagactcggagtgcttccaaagtgc 120
Db 192 AGTGCAGCTTCCCTTGGCGGTGGTGAAGAGACTCGGGAGTCTGCTTCCAAAGTGCC 251
QY 121 cgcgtgagtgagctctcaccctcagcagcaaatgagcctcttcgggtcttccctgctg 180
Db 252 CGCCGTGAGTGAGCTCTCACCCTCAGCCCAATGAGCTCTTCGGGCTTCTCTGCTG 311
QY 181 acatctgcctggccggccagagacaggggactcagcggaatcccaacctgagtagtaa 240
Db 312 ACATCTGCCCTGGCCGCCAGAGACAGAGGAGCTCAGGCGGAATCCAACCTGAGTAGTAAA 371
QY 241 ttccagttttccagcaacaaggaacagagcagagtcacagatcctcagatgagagaatt 300
Db 372 TTCCAGTTTTTCCAGCAACAAGGAACAGACGGAGTACAAGATCCTCAGCATGAGAAATT 431
QY 301 attactgtctactaatggaagtattcacagcccaaggttttctctactatcttcaaga 360
Db 432 ATTACTGTCTACTAATGGAAGTATTCAGCCCCAAGGTTTCTCTCATATCTTATCCAAGA 491
QY 361 aatacggcttggtaggagatttagcagtagagagaaatgatggatacaacttacg 420
Db 492 AATACGGCTTGGTATGGAGATTAGTAGCAGTAGAGGAAAAATGATGATACAACTTACG 551
QY 421 ttgtatgaagatttgggttgaagaccagagaagatgcacatatgcgaagtatgatttga 480
Db 552 TTTGATGAAAGATTTGGGCTTGAAGACCCAGAGATGACATATGCAAGTATGATTGTTGA 611

QY 481 gaagttgaggaaccagtgatggaactatattagggcgctggtgtggttctggttactgtga 540
Db 612 GAAGTTGAGGAACCCAGTGATGGAACATATATTAGGGCGCTGCTGGTTCGTGACTGTA 671
QY 541 ccaggaaacagatttcaagaagaatacaattaggataagattgtatctgatgaatat 600
Db 672 CCAGGAAACAGATTTCTAAGGAAATCAAAATTAGGATAAGATTGTATCTGATGAATAT 731
QY 601 ttctctctgaaccaggggttctgcatccactacaacttgcattgctgcacacaattcacagaa 660
Db 732 TTTCCCTCTGAACACGAGGTTCTGCATCCACTACACATATTCATGCGCCACATTCACAGAA 791
QY 661 gctgtgagtccttcagtgctaccoccttcagcttgcactggaccctgcttaataatgct 720
Db 792 GCTGTGAGTCCTTCAGTGCTACCCCTTCAGCTTTGCCACTGGACCTGCTTAAATAATGCT 851
QY 721 ataaactgcttttagtaaccttggagaaccttattogatactattgtgaaccagagagatggcag 780
Db 852 ATAACCTGCCPTTATAGTACCTTGGAGAGACCTTATTCGATATCTTGAACACAGAGATGGCAG 911
QY 781 ttgacttagaagatctatatagcccaacttgcaacttcttggcaaggcttttggtttt 840
Db 912 TTGACTTAGAAGATCTATATAGCCCAACTTGGCAACTTCTTGCAAGGCTTTTGTGTTTT 971
QY 841 ggaagaaatccagagtggttgatctgaaaccttctaacagaggaggtgaagattatacagc 900
Db 972 GGAAGAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAAGATTATACAGC 1031
QY 901 tgcaacctgtaacttctcagtgctcataagggagaactaaagaaacagacgataccatt 960
Db 1032 TGCACACCTCGTAACCTTCAGTGCTCCATAGGGAAGAACTTAAAGAGAAACCGATACCAT 1091
QY 961 ttctgcccaggttctcctggttaaacgctgtgtgggaactgctgtgtgtctccac 1020
Db 1092 TTTGCGCCAGGTGTCTCCGTTTAAACGCTGTGGGAACTGTGCTGTGTGCTCCAC 1151
QY 1021 aattgcaatgaatgtcaatgtgtcccaagcaaatgtactaaataataccagaggtcctt 1080
Db 1152 AATTGCAATGAATGTCAATGTGTCCCAAGCAAAAGTTACTAAAAAATACCACGAGTCTT 1211
QY 1081 cagttgagacaaagaccggtgtcaggggattgcaaaaactcaactcaacgcagtgccctg 1140
Db 1212 CAGTTGAGACCAAGACCGGTGTGAGGGGATTGACAAAATCACATCACCGAGTGCCCTG 1271
QY 1141 gacacacatagcagtgctgctgtgagagggagacagagataagccacatcac 1200
Db 1272 GACACATGAGAGTGTGACTGTGTGAGAGGGAGCACAGAGGATAGCCGATCAC 1331
QY 1201 caccagcagctctgcccagagctgtgcagtllygvcagtggtgattcttattagagaa 1260
Db 1332 CACCAGCAGCTCTGCCACAGCTGTGCAGT- ---GCAGTGGCTGATCTATTAGAGAAC 1387
QY 1261 gtatgcgttatccatccatctcagttgttggtttccaaagacatttccattcagg 1320
Db 1388 GTATGCGTTATCTCCATCTTAATCTCAGTGTGTTTGGCTTCAAGGACCTTTCATCTCAGG 1447
QY 1321 attacagtgcaattctgaagagagagacatcaacaagaattagagttgtgcaacagctc 1380
Db 1448 ATTTACAGTGCATCTCTGAAGAGAGAGACATCAAAACAGAAATTAGAGATTGTCAACAGCTC 1507
QY 1381 ttttgagagagagcctaaagacagagagaaaaggtcttccatcgtgaaagaaaataaa 1440
Db 1508 TTTTGAGAGAGAGGCTTAAGACAGAGAGAAAAGGCTTCAATCTGGAAGAAAATTA 1567
QY 1441 tgttgtattaaatagatcacagcagtagtttccagagttaccatgtacgtatttccactagct 1500
Db 1568 TGTGTGATTAAATAGATCACAGCTAGTTTCCAGAGTTTACCATGTACGTATTCCACTAGCT 1627
QY 1501 ggggtctgtatttcagttcttctogatacggcttagggtaattgtcagtacagggaaataaac 1560
Db 1628 GGGTCTGTATTTCAGTCTTTCATACGGCTTAGGGTAAATGTTCAGTACAGGAAAAAAC 1687
QY 1561 tftgcaagtgagcacctgattccgttgccttgccttaactcaagctccatgctcctgggc 1620

Db 1688 TTGTGAAGTGAACACCTGATTCGGTTGCCTTAACTAAAGCTCCATGCTCGTGGC 1747
QY 1621 ctaaaatcgataaaatcgatgttttttttttttttttttttttttttttttttttttttt 1680
Db 1748 CTAATAATCGTAAATAATCTGGA- TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 1806
QY 1681 accagaacattctatgtactacaacactggttttttaaaaggaactatgttgcctatgaat 1740
Db 1807 ACCAGAACATTTCTATGTACTACAAACCTGGTTTAAAGGAAGCACTATGTTGCTATGAAT 1866
QY 1741 taaacttgtctgctgctgtagga 1764
Db 1867 TAAACTTGTCTATGCTGATAGGA 1890
RESULT 4
AF244813 2152 bp mRNA PRI 01-JUL-2000
LOCUS Homo sapiens platelet-derived growth factor C mRNA, complete cds.
DEFINITION AF244813
ACCESSION AF244813
VERSION AF244813.1 GI:8886883
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2152)
AUTHORS Li,X., Ponten,A., Aase,K., Karlsson,L., Abramsson,A., Utela,M.,
Backstrom,G., Hellstrom,M., Bostrom,H., Li,H., Soriano,P.,
Betscholtz,C., Heldin,C.-H., Allitalo,K., Ostman,A. and Eriksson,U.
PDGF-C is a novel protease-activated ligand for the PDGF alpha
receptor
JOURNAL Nat. Cell Biol. (2000) In press
REFERENCE 2 (bases 1 to 2152)
AUTHORS Eriksson,U., Aase,K., Li,X. and Ponten,A.
TITLE Direct Submission
JOURNAL Submitted (14-MAR-2000) Ludwig Institute for Cancer Research,
Nobels vag 3 P.O.Box 240, Stockholm S-171 77, Sweden
FEATURES
location/Qualifiers
1..2152
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="lung"
/dev_stage="fetal"
189..1226
/function="ligand and agonist for platelet-derived growth
factor receptor alpha (PDGFalpha)"
/note="PDGF-C; member of the PDGF/VEGF growth factor
family; contains N-terminal CUB domain and the C-terminal
PDGF/VEGF-like domain"
/codon_start=1
/product="platelet-derived growth factor C"
/protein_id="AAF80597.1"
/db_xref="GI:8886884"
/translation="MSLRGLLVTSALAGORGTQAESNLSSKFQFSNKGNGVQDP
QHERLIIVSTNGSIHSRPFHTYPRNTVLVRLVAENVWVQITDFRFGLEPEDD
ICKYDFVEVEEPSDGLILGRWCVGSGTGVKQISKQNIIRFVSDEYSPSPGFCIH
NIVMPQTEAVSPSVLPALPLDLNNATFSTLEDLIRLEPERWQLEDLYRP
TWLLGKRAVFGKRSVVDNLNLTVEEVLRYCTPRNFVSIRLEPKRDTTFWPGCLL
VKRCGNGACCLHNCNECCVPSVKYKHYEVLQRPKTVGRGLKSLTDVALEHHEE
CDCVCRGSTGG"
BASE COUNT 621 a 424 c 476 g 630 t 1 others
ORIGIN
Query Match 96.4%; Score 1701; DB 89; Length 2152;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1752; Conservative 1; Mismatches 6; Indels 8; Gaps 4;
QY 1 attatgtggaactaccctgcgattctctgctgcagagcagctgcgctccacccc 60
Db 33 ATTATGTGAAGAACTACCTTCGATTTCTGTCTGCCAGACAGCTCGGCGCTTCCACCCC 92

/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="embryo"
/tissue_type="brain"
327..1364
/gene="hscdgp"
CDS
327..1364
/gene="hscdgp"
/codon_start=1
/product="spinal cord-derived growth factor"
/protein_id="BA03266.1"
/db_xref="GI:932294"
/translation="MSLFGLLITSALAGORQGTQAESNLSKQFSSNBQNVQDP
QHERITTSVGHSPFPHYPRNTLVLRVAVVEENWMIQTFDFRFGLEDPEDD
ICKYDFVEPEPDSGTILGRWCSGTVPKQISKGNOIRIFVSDVEYFPSPGFCIH
NIMPOFTSAVSPVLPSPSALPDLLNNAITAFSTLEDLIRYLEPERWOLDEDLYRP
TWMLGKAFVGRKSRVDLNLITVEVRLYSCYPRNFSVSTRLEKLTDTDFWGCLL
VRKGGNCACCLHNCNEQCVPKVKYKHYEVQLRPKTVGRGLHKSUTDVALEHHEE
CDCVCRGSGG"

BASE COUNT 501 a 412 c 424 g 477 t 3 others
ORIGIN

Query Match 89.2%; Score 1574; DB 85; Length 1817;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1629; Conservative 0; Mismatches 10; Indels 9; Gaps 4;

QY 1 attatgtggaactaccctcgattctctgctgcagagcagctcgcgcttccacccc 60
Db 174 ATATGTGGAAACFACCCTCGGATCTCTGCTGCCAGAGCGCTCGCGCTTCCACCCC 233
QY 61 agtcagccttccctggcgtggtgaaagactcggagtcgctgtctccaaagtgcc 120
Db 234 AGTCAGCCTTCCCTGGCGTGTGAAAGAGACTCGGGAGTCGCTGCTTCCAAAGTGC 293
QY 121 cgcgctgagtgagctctcaccctcagtcagccaaatgagctctctggctctctctg 180
Db 294 CGCGCTGAGTGAGCTCTACCCAGTCAGCCAAATGAGCCTCTTCGGGCTTCTCGTGTG 353
QY 181 acatctccctggcggccagagacaggggactcaggcggaatccaaactgagtagtaa 240
Db 354 ACATCTGCCCTGGCGCCAGAGACAGGGAGTCTCAGCGGAATCCAACTGAGTAGTAAA 413
QY 241 ttccagtttccagcaagaagaacagcagagtgacagatcctcagcatgagagaatt 300
Db 414 TTCCAGTTTCCAGCAACAGAACAGACGAGTACAGATCCTCACCATGAGAGATT 473
QY 301 attactgtctactaatggaagtattccacagcccaaggtttcctcactatccaaga 360
Db 474 ATTACTGTCTACTAATGGAAGTATTACACGCCCAAGTTTCCTCATACTATCCAAGA 533
QY 361 aatacgtcttgtagagattagtagcagtagaggaataatgatgatacaacttacg 420
Db 534 AATACGGTCTGGTATGAGGATTTAGTACAGTAGAGGAAATGATGGATACAACTTACG 593
QY 421 ttgatgaagatttggcttgaagaccagaaagatcacatgacatgattgatttga 480
Db 594 TTGATGAAGATTGGGCTTGAAGACCCAGAACATGACATGACATGATGATTTGTA 653
QY 481 gaagttgaggaaccagtgatggaactatataggcgctggtggtggtctggtactga 540
Db 654 GAAGTTGAGGAACCCAGTGATGGAACATATTAGGGCGCTGCTGGTCTGGTACTGTA 713
QY 541 ccaggaacaacagattcttaaggaaatcaaataggataagattgttatctgatgaatat 600
Db 714 CCAGGAACAACAGATTCTTAAAGGAAATCAAAATTAGGATAAGATTGTATCTGATGAATAT 773
QY 601 ttctctctaaaccaggttctgcatccactacaacattgtcatgccaaattcacagaa 660
Db 774 TTTCCTTCTGAACACGAGGTCTGTCATCCACTACAAACATTTGTCATGCCAAATTCACAGAA 833
QY 661 gctgtgagctcttcagtgctaccctctcagcttctgacactggacctgcttaataatgct 720
|||||

Db 834 GCTGTAGTCCCTTCAGTGTGTACCCCTTCAGCTTTTGCCACTTGACCTGCTTAATATGCT 893
QY 721 ataaatgccttttagtaccttgaagaccttattcatatcttgaaccagagagatggag 780
Db 894 ATAACTGCCCTTTAGTACCTTGGAAAGACCTTATTCGATATCTTGAACCCAGAGATGGCAG 953
QY 781 ttggacttagaagatctatatagggccaacttggccaacttcttggcaagccttttgtttt 840
Db 954 TTGGACTTGAAGATCTATATAGGCCAACTTGCAACTTCTTGGAAGGCTTTTGTGTTTT 1013
QY 841 ggaagaaatccagagtggtgatctgaacctcttaacagagagaggaagatatatacagc 900
Db 1014 GGAAGAAATCCAGAGTGTGTGATCTGAACCTTCAACAGAGAGGTAAGATTATACAGC 1073
QY 901 tqcacacctcgttaactctcaatgtccataaggggaagaaactaaagagaaccatcaccatt 960
Db 1074 TGCACACCTCGTAACCTTCTCAGTGTCCATPAGGGAGAAACTTAAGAGAACCCGATACCAT 1133
QY 961 ttctggccaggttctctctctggtttaaagcgtgtggtggaaactgtgcttgttctccac 1020
Db 1134 TTCTGCCAGGTTGTCTCTCTGTTAAACGCTGTGTGGTGGAACTGTGCCGTGTTCTCCAC 1193
QY 1021 aattgcaatgaatgtcaatgtgtcccaagcaagttactataaaataccacagagtcctt 1080
Db 1194 AATTGCAATGAATGTCAATGTGTCTCCAGCAAAAGTTACTTAAAAAATACCACGAGTCTCT 1253
QY 1081 cagttgagaccaaaacacgggtgtcaggggattgcacaaactcaactcacgacgtgccctg 1140
Db 1254 CAGTTGAGACCAAAAGACCGGTGTGAGGGATTGACAAATCACTCACCAGCTGGGCCCTG 1313
QY 1141 gagcaccatgagaggtgtgactgtgtgcagagggagcacagggagtagcgcgcatcac 1200
Db 1314 GAGCACCATGAGGAGTGTGACTGTGTGTCAGAGGGAGCACAGGAGTAGCGCATCAC 1373
QY 1201 caccgagcgtcttgcacagagctgtgcagtllylncagtggtgattctattagagaac 1260
Db 1374 CACCAGCAGCTTTGCCAAAGCTGTGCAGT---GCATGGCTGATTCTATTAAAGAAC 1429
QY 1261 gtatcggtattctcatcttaacttaactcagttgttcttcaaggaccttctcagg 1320
Db 1430 GTATCGGTTATCTCCATCCTTAATCTAGTTGTTTGTCTTCAAGGACCTTTTCATCTCAGG 1489
QY 1321 attcagtgctattctgaaagaggagacacatacaagaattaggagttgtgcaacagctc 1380
Db 1490 ATTTACAGTGCATCTTGAAGAGAGGACATCAACAGAGATTAGGAGTTGTGCAACAGCTC 1549
QY 1381 ttttgagggagggccttaagagcagagagaaagggttctcaatcgttggaagaaaaattaaa 1440
Db 1550 TTTTGAGAGGAGGCGCTTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATAAA 1609
QY 1441 tttgtattaaatagatcacacagctagtttcaagattaccatg-tacgtattccactagc 1499
Db 1610 TGTGTATTAATAATAGATCACCGCTAGTTTTCAGAGTTACCATGTTACGTATTCTCCACTAGC 1669
QY 1500 tgggtctgtattcagttcttctcagatcagcgttagggttaat-gtcagtcacaggaaaaa 1558
Db 1670 TGGGTTCTGTATTTCAGTCTTTTCGNATACAGCTTAGGGTAATAGTCAGTACAGGCAAAA 1729
QY 1559 actgtg---caagtgagcacctgattccgcttgccttgccttaactcctaaagctcccatgtcc 1615
Db 1730 ATACTGTGCAGAGTATCACCTGATTCCTGCTTGCCTTGGCTTAACTCTTAAGTCTCATGTCC 1789
QY 1616 tgggcttaaaatcgtataaaaatcctggat 1643
Db 1790 TGGGCTTAAATCGTATAAAATCTGGAT 1817

RESULT 6
AX027935
LOCUS AX027935 1035 bp DNA
DEFINITION Sequence 3 from Patent WO0037641.
ACCESSION AX027935
VERSION AX027935.1 GI:10188752

| | |
|----------|--|
| KEYWORDS | human. |
| SOURCE | ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1035) Gordon,R.D., Dijkmans,J.J. ¹ Sprengel,J.J., Yon,J.R., Xu,J., Gosiowska,A. and Dhanaraj,S.N. Patent: WO 0037641-A 3 29-JUN-2000: Vascular endothelial growth factor-X GORDON ROBERT DOUGLAS (BE) ; DIJKMANS JOSTENA JOHANNA HUBER (BE) ; JANSSEN PHARMACEUTICA NV (BE) ; SPRENGEL JORG JURGEN (BE) ; YON JEFFREY ROLAND (BE) ; XU JEAN (US) ; GOSIEWSKA ANNA (US) ; DHANARAJ SRIDEVI NAIDU (US) FEATURES source Location/Qualifiers 1..1035 /organism="Homo sapiens" /db_xref="taxon:9606" BASE COUNT 301 a 213 c 251 g 270 t ORIGIN Query Match 58.7%; Score 1035; DB 9; Length 1035; Best Local Similarity 100.0%; Pred. No. 2.9e-266; Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 154 atgagccttcctggccttcctcgtcacatctccctgccgcccagacacagcggact 213 Db 1 ATGAGCCTCTTCGGCGCTTCCTCGTCAATCTGCCCTGGCCGCCAGACAGGGGACT 60 QY 214 caggcgggaatccaacctgagtagtaaatctccagttctccagcaacaaggaaacacgga 273 Db 61 CAGSCGGAATCCAACCTGAGTAGTAATTCAGGTTTCCAGCAACAAGGAACGGA 120 QY 274 gtacaagatcctcagcatgagagaattattactgtctactaatgaagtattcaacgc 333 Db 121 GTACAAGATCCTCAGCATGAGAGAATTATTACTGTGCTACTAATGGAAGTATTTCACAG 180 QY 334 ccaaggttctctatacttaccagaataacgctctgtatgagattagatagacgta 393 Db 181 CCAAGGTTCTCTATACTATCCAAAGAAATACGGCTCTGGTATGGAGATTAGTAGCAGTA 240 QY 394 gaggaataatgatggatacaactcacgtttgatgaaagattggcgttgaaaccccgaa 453 Db 241 GAGGAAATGTATGGATACAACTTACGTTTGATCAAGATTTCGGCTTGAAGACCCAGAA 300 QY 454 gatgacatgcaagtgatgatttgtagaagttgaggaaacccagtgatggaactatatta 513 Db 301 GATGACATATGCAAGTATGATTTGTAGAGTTGAGAGTTGAGGAACCCAGTGATGGAACATATATTA 360 QY 514 gggcgctgggtggttcctgctactgtaccagaaaacagattcttaagaaatcaaat 573 Db 361 GGGCGCTGGGTGGTCTGGTACTGTACCAGAAAACAGATTCTTAAGGAATTCAAAT 420 QY 574 aggaagaagttgtatctgataaattttctctctgaacccagggttctgcataccactac 633 Db 421 AGGATAAGATTGTATCTGATGAATATTTCTCTTCAACACGGGTTCTGCATCCACTAC 480 QY 634 acaattgcatagcacaattcaacagagctgtgagtccttcagtcacccccctcagct 693 Db 481 AACATTGTCATGCCACAAATTCACAGAAGCTGTAGTGCTTCAGTGCCTACCCCTTCAGCT 540 QY 694 ttgcactggacgtcgttaaatgctataactgctcttagtaccttggagaaccttatt 753 Db 541 TTGCCACTGGACCTGCTTAATAATGCTATAACTGCCCTTTAGTACCTTTGGAAGACCTTATT 600 QY 754 cgatatcttgaacacagagatggcagttggacttagaagattctatataggccaacttg 813 Db 601 CGATATCTTGACACAGAGAGTGCAGTTGGACTTGAAGATCTATATAGCCCACTGG 660 QY 814 caactcttgcaaggctttgttttggagaataatccagagtggttgatctgaacctt 873 Db 661 CAACTCTTCGGCAGGCTTTGTTTTTGGAGAAATTCAGAGTGGTGGATCTGAACCTT 720 |

| | | | | | | | | |
|-----------------------|---|---|-------|---------------------|-------------|---------------------|-------|------|
| Qy | 874 | ctaacagaggaggtaagattatacagctgcacaccctcgtaaacttctcagtgctccataagg | | | | | | 933 |
| Db | 721 | CTACAGAGGAGGTAAAGATTATACAGCTGCACACCCTCGTAACCTTCTCAGTGTCATAAAG | | | | | | 780 |
| Qy | 934 | gaagaactaaagaaacccgatacaccaatttttcggccagattgtctcctggttaaacgctgt | | | | | | 993 |
| Db | 781 | GAAGAACTAAAGAAACCCGATACCAATTTCCTGGCCAGGTGTCTCCTGGTTAAACGCCGTG | | | | | | 840 |
| Qy | 994 | ggtgggaactgtgccctgttgtctccacaattgccaatgaatgtccaatgtgtccccaaagcaaa | | | | | | 1053 |
| Db | 841 | GGTGGRACTGTGCCTGTGTGCTCCACAATTCGAATGATGATCATGTGTCCCAGACAAA | | | | | | 900 |
| Qy | 1054 | gttactaaaaataaccacagaggtccttcagttagaccaaaagacgggtgtcaggggattg | | | | | | 1113 |
| Db | 901 | GTTACTAAAAAATACCACGAGTCTTCAGTTGAGACCAAAGACCGGTGTCCAGGGGATTG | | | | | | 960 |
| Qy | 1114 | cacaatacactcaccgacgtggccctgagaccacatgaggatgtgaactgtgtgtgcaga | | | | | | 1173 |
| Db | 961 | CACAAATCACTCACGAGCGTGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGCAGA | | | | | | 1020 |
| Qy | 1174 | gggagcacaggagga | | | | | | 1188 |
| Db | 1021 | GGGACACAGGAGGA | | | | | | 1035 |
| <hr/> | | | | | | | | |
| RESULT | | 7 | | | | | | |
| AX028032 | | | | | | | | |
| LOCUS | AX028032 | 1035 bp | DNA | PRI | 16-SEP-2000 | | | |
| DEFINITION | Sequence 3 from Patent WO0037641. | | | | | | | |
| ACCESSION | AX028032 | | | | | | | |
| VERSION | AX028032.1 GI:10188844 | | | | | | | |
| KEYWORDS | human. | | | | | | | |
| SOURCE | ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1035) Gordon,R.D., Dijkmans,J.J., Sprengel,J.J., Yon,J.R., Xu,J., Gosiewska,A. and Dhanaraj,S.N. Vascular endothelial growth factor-x Patent: WO 0037641-A 29-JUN-2000; GORDON ROBERT DOUGLAS (BE) ; DIJKMANS JOSIENA JOHANNA HUBER (BE) ; JANSSEN PHARMACEUTICA NV (BE) ; SPRENGEL JORG JURGEN (BE) ; YON JEFFREY ROLAND (BE) ; XU JEAN (US) ; GOSIEWSKA ANNA (US) ; DHANARAJ SRIDEVI NAIDU (US) | | | | | | | |
| FEATURES | Location/Qualifiers source 1..1035 /organism="Homo sapiens" /db_xref="taxon:9606" | | | | | | | |
| BASE COUNT | 301 a | | 213 c | | 251 g | | 270 t | |
| ORIGIN | | | | | | | | |
| <hr/> | | | | | | | | |
| Query Match | | 58.7% | | Score 1035; | | DB 91; Length 1035; | | |
| Best Local Similarity | | 100.0%; | | Pred. No. 2.9e+266; | | | | |
| Matches 1035; | | Conservative 0; | | Mismatches 0; | | Indels 0; Gaps | | |
| Qy | 154 | atgagcctcttcggggtctcctcgtgcatcttgcctcgccggccgagacagagcaggagact | | | | | | 213 |
| Db | 1 | ATGAGCCTCTTCGGGCTTCTCCTGCTGACATCTGCCCTGGCGGCAGACAGACGGGACT | | | | | | 60 |
| Qy | 214 | caggcgaatccaacctgagtagtaaatccagtttccagcaacaaggaacagaacgga | | | | | | 273 |
| Db | 61 | CAGCGGAATCCAACCTGAGTAGTAAATTCAGTTTTCAGCAACAAAGGAACAGAACGGA | | | | | | 120 |
| Qy | 274 | gtacaagatcctcagcatgagagaattattactgtgttacttaatgggaagtattcacagc | | | | | | 333 |
| Db | 121 | GTACAGATCCTCAGCATGAGAGAATTATTAATTGTTGTTCTACTTAATGGAAGTATTCACAGC | | | | | | 180 |
| Qy | 334 | ccaaggttctctcatacttatccaagaataacggtcttggtagtagattagtagcaqta | | | | | | 393 |
| Db | 181 | CCAAGGTTCTCTACATATATCAAGAANAATACGGTCTTGGTATGAGATTAGTAGCAGTA | | | | | | 240 |

QY 394 gaggaatgtatggatacaacttaacgtttgatgaagatttggcttgaagaccagaa 453
Db 241 GAGGAAATGTATGGATACAACTTACGTTTGTATGAAAGATTGGCTTGAAGACCCAGAA 300
QY 454 gatgaatatgaagatgatattttagaagttgaggaacccagtgatggaactatatta 513
Db 301 GATGACATATGCAAGTATGATTTTGTAGAAGTTGAGGAACCCAGTGATGGAACATATATTA 360
QY 514 ggcgcgtggtgttctgtactgtaccaggaacacagatttctaaggaatacaaat 573
Db 361 GGGCGCTGGTGTGCTGTCTGTACTGTACCAAGAAACAGATTTCTAAGGAAATCAAAAT 420
QY 574 agataagatttgcataatatttccctctgaacagaggttctgtcatccactac 633
Db 421 AGGATGAAGATTGTATCTGTATGAATATTTCTTCTGAACAGAGTCTTGCAATCACTAC 480
QY 634 aacattgtcatgccaattcaagaagctgtgagtccttaagtgtaccaccccttcagct 693
Db 481 AACATTGTCTATGCCACAAATTCACAGAAGCTGTGAGTCTCTCAGTGTACCCCTTCAGCT 540
QY 694 ttccactgacctgtctaataatgctataactgccttttagtaccttggagacattatt 753
Db 541 TTGCCACTGGACCTGCTTAATATATGCTTAATACCTGCTTTAGTACCTTGGAAAGACCTTAAT 600
QY 754 cgatatcttgaaccagagatggcagttggacttagaagatctatatagggccaacttgg 813
Db 601 CGATATCTTCAACACAGAGATGGCAGTTGGACTTTAGAAGATCTATATAGGCCAACTTGG 660
QY 814 caacttcttggcaggttcttggtttttgggaagaaatccagagtggtggaactgaacctt 873
Db 661 CAACCTTCTTGGCAAGGCTTTTGTGTTTGGGAAGAAATCCAGAGTGTGGATCTGAACCTT 720
QY 874 ctacagagagatgaagattatacagctcacacctgtaactctcaactctcaagtccataagg 933
Db 721 CTAAACAGAGAGGTAAGATTATACAGCTGCACACCTCGTAACCTTCTCAGTGTCCATAAGG 780
QY 934 gaagaactaaagagaacccagctacatttctggccaggtgtctcctggtttaaaccgtgt 993
Db 781 GAAGAATAAAGAGAACCGATACATTTTCTGGCCAGGTTGTCCTCTTAAACGCTGT 840
QY 994 ggtgggaactgtgctgttctctccacaattgcaatgaatgtcaatgtgtcccaagcaaa 1053
Db 841 GGTGGGAACCTGTGCTTCTCTCCACAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 900
QY 1054 gttactaaaaataaccagaggtccttcagttgagacaaagacccggtgtcaggggattg 1113
Db 901 GTTACTAAAAAATACCACGAGGCTTTCAGTTGAGACAAAGACCGGTGTCCAGGGGATTG 960
QY 1114 cacaataactcaccgacgtggccctggagcaccatgagagtggtgactgtgtgtgcaga 1173
Db 961 CACAAATCATCACCGACGTGGCCCTGAGCACCATGAGGAGTGTGACTGTGTGTGCAGA 1020
QY 1174 gggagcacaggagga 1188
Db 1021 GGGAGCACAGGAGGA 1035

RESULT 8
AX044520
LOCUS AX044520 3571 bp DNA PAT 24-NOV-2000
DEFINITION Sequence 34 from Patent WO0066736.
ACCESSION AX044520
VERSION AX044520.1 GI:11343375
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3571)
AUTHORS Gilbert,T., Hart,C.E., Sheppard,P.O. and Gilbertson,D.G.
TITLE Growth factor homolog zvegfa
JOURNAL Patent: WO 0066736-A 34 09-NOV-2000;

ZymoGenetics, Inc. (US)
Location/Qualifiers
1. 3571
/organism="Mus musculus"
/db_xref="taxon:10090"
CDS
1049..2086
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC17171.1"
/db_xref="GI:11343376"
/translation="MLLGLLLLSALAGQTGTAEENLSKQLQLSSDKQNGVQDD
RHERVITISNGSHSPKPHPTPRNMYLVRLVAVDENVRIQTFDRFLEDPEDD
ICKYDFEVEEPPSGVLRGWCSTGKQTSKGNHIRIRFVSEYFPPGFCIH
SIIMPQVETTSPLVPSLSLLDNNATVAFSELEIRFVLPDRQVLDLSLYKP
TWOLGKAFLYGKSKVNLNLLKEVKLYSCTPRNFVSIREELKRTDTTFPGCLL
VKRCGNCACCLHNCQCQVPRKVTYKHYEVLQLRPKTKYKGLHKSITDVALEHHE
CDCVCRGNAGG"
BASE COUNT 876 a 935 c 875 g 885 t
ORIGIN
Query Match 50.1%; Score 883.2; DB 9; Length 3571;
Best Local Similarity 83.7%; Pred. No. 1.5e-225;
Matches 1014; Conservative 0; Mismatches 193; Indels 5; Gaps 1;
QY 1 attatgtgaaactaccctgcgattctctgtgcagagcagctgcgcgttcacccc 60
Db 901 ATTATGTGAAACTACCTGCGATTCTCTGTGTCAGAGCCGCCGCGCTTCCACCGC 960
QY 61 agtcagacctccctggcgtgtggtgaaagagactcgggagtcgctgtctccaaagtgc 120
Db 961 AGCGCAGCGCTTCCCGCGCTGGGTGA-----GCCTTGGAGTCTGCTGCCAGTGC 1015
QY 121 cgcctgagtgagctctcaccctcagcgaataatgagcctcttctgggtctctcgtg 180
Db 1016 CGCCCGAGTGAAGCTCTGCGCCAGTCAGCCAAATGCTCTCTCGGCTCTCTGCTG 1075
QY 181 acatctgccttgcgcgcgcagagagcagggactcagggcgaatccaaacctgagtgaaa 240
Db 1076 ACATCTGCTGCGCGCGCCAAAGAACGGGACTCGGGCTGAGTCCAACTGAGCAGCA 1135
QY 241 ttccagtttccagcaacaaagaaacagagagactcctcagatcctcagcatgagagaat 300
Db 1136 TTGACGCTCTCCAGCGACAAAGAACAGACGAGTGAAGATCCCGGCTCAGAGAGTT 1195
QY 301 attactgtctactaatggaagtattcacagcccaagtttccctcacttactatccaga 360
Db 1196 GTCACTATATCTGTTAATGGGAGCATCCACAGCCCGAAGTTTCTCTCATATACCAAGA 1255
QY 361 aatcgcgtcttgatgagatttagcagtagaggaataatgtagatacaacttaacg 420
Db 1256 AATATGTGCTGCTGGTGGAGATTAGTTGAGTAGATGAAATGTGCGGATCAGCTGACA 1315
QY 421 ttgatgaagatttggccttgaagaccagagagatgacatatgcaagtattgatttga 480
Db 1316 TTTGATGAGAGATTGGCTGGAAGATCCAGAAAGACGATATATGCAAGTATGATTTGTA 1375
QY 481 gaagttgaggaacccagtgatggaactatttagggcgtggtggttctgtgactgta 540
Db 1376 GAAAGTTGAGGACCCAGTGATGGAAGTGTTTAGGACGCTGCTGTTCTGGGACTGTG 1435
QY 541 ccaggaacacagattcttaaggaatacaaataggaataagattgtatctgatgaatat 600
Db 1436 CCAGGAAAGCAGACTTCTAAAGAAATATATATGAGAAATGATGATGATGATGATGAT 1495
QY 601 ttctcttgaaacaggttctgcactacacattgcatcccaattccacagaa 660
Db 1496 TTTTCATCTGAACCCGGATTCTGCATPCCACTACATATATATCATGCCACAGTACAGAA 1555
QY 661 gctgagctcttcagtgctaccccttcagcttgcagctgagcctgacctgacttaataatgt 720
Db 1556 ACCACGAGTCTCTCGTGTGTGCCCTTCATCTTTGTCATTTGGAGACTGCTGCAACATGCT 1615

QY 721 ataactgccttagtaccttggaagacccttattcgatatctcttgaaaccagagagatggcag 780
Db 1616 GTGACTGCTTCAGTACCTTGAAGAGCTGATTCGGTACTAGAGCAGCATCGATGCGCAG 1675
QY 781 ttggacttagaagatctatataggccaacttgcaactctctggcaagcttttggttttt 840
Db 1676 GTGGACTTGGCAGCGCTCTACAGCCACATGCGACGCTTTGGCAAGGCTTTCCCTGTAT 1735
QY 841 ggaagaaatccagagtggtgagctgaaccttcttaacagagagagtgtaagattatcacgc 900
Db 1736 GGGAAAAAAGCAAGTGGTGAATCTGAATCTCCTCAAGGAGAGGTAAACCTCTACAGC 1795
QY 901 tgcacacctcgtaactctcagtgccataagggaagaaactaaagaaacccgataccatt 960
Db 1796 TGCACACCCCGGAACCTTCTCAGTGTCCATACGGAAGAGCTAAAGAGGACAGATACCATA 1855
QY 961 tctggccaggtgtcctcgtgtaaacgctgtgtggaactgtcctgtgtgttctccac 1020
Db 1856 TTCTGGCCAGGTTGCTCTCTGCTCAAGCGCTGTGGAGGAATTTGCGCTGTGCTCCAT 1915
QY 1021 aattgcaatgaatgtcaatgtgtcccaagcaaaagtgtactaaaaataaccacagaggtcctt 1080
Db 1916 AATTGCAATGAATGTCAGTGTGTCACCGCTAAAGTTACAAAAAAGTACCATGAGGTCCCT 1975
QY 1081 cagttgagacaaagaccggtgtcaggggattgcacaaatcactcaccgacgtggccctg 1140
Db 1976 CAGTTGAGACCAAAACCTGGAGCTCAAGGGATTGCATAAGTCACTCACTGATGTGGCTCTG 2035
QY 1141 gagcaccatgagaggtgtgactgtgtgacagagagcacagagagatagccgacacac 1200
Db 2036 GAACACCAGAGGAATGTGCTGTGTGTAGAGGAACCGCAGGAGGTAACTGCGACGCT 2095
QY 1201 caccagcagctc 1212
Db 2096 TCGTAGCAGCAC 2107
RESULT 9
AF117608 2692 bp mRNA ROD 02-JAN-2000
LOCUS Mus musculus fallotein mRNA, complete cds.
ACCESSION AF117608
VERSION AF117608.1 GI:6652867
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2692)
AUTHORS Tsai,Y.-J., Lee,R.K.-K., Chen,Y.-H., Lin,S.-P. and Cheng,W.T.-K.
TITLE cDNA cloning of fallotein from mouse ovary
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2692)
AUTHORS Tsai,Y.-J., Lee,R.K.-K., Chen,Y.-H., Lin,S.-P. and Cheng,W.T.-K.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-1999) Medical Research, Mackay Memorial Hospital,
45 Min Sheng Road, Tamshui, Taipei 25115, Taiwan
FEATURES
source 1..2692
/organism="Mus musculus"
/db_xref="taxon:10090"
/sex="female"
/tissue_type="ovary"
/dev_stage="adult"
198..1235
/note="putative secretory protein"
/codon_start=1
/evidence=not_experimental
/product="fallotein"
/protein_id="AAF22516.1"
/db_xref="GI:6652868"
/translation="MLLGLLLLSALAGQRTGPAESNLSSKILQLSDKQNGVQDP
RHERVVTISGNISHSFPKFTYPRNMLVWRVAVDENVRILQTFDFRFLGLEPDD

polyA_signal 754 a 576 c 617 g 745 t
BASE COUNT 2662..2667
ORIGIN
ICKYDFVEVEEPSDGSVLGRWCOSGTVPGKQTSKGNHIRFVSDYFFSEPGFCIH
STIMPQTFETSPSPSSLDLNNAVTAFSTLEELIRYLEPDRWVDLSLYKP
TWOLLKAFLYGKSKVNLNLLKEEVLYSCTPRNFVSIRLELRKTDITFWPGCLL
VKRCGNCACCLHNCNECCQPKVTKKYHEVLQLRPKTKVGLHLSLTDVALEHHEE
CDCVCRNAGG"
2662..2667
Query Match 50.0%; Score 881.6; DB 94; Length 2692;
Best Local Similarity 83.6%; Pred. No. 4e-225;
Matches 1013; Conservative 0; Mismatches 194; Indels 5; Gaps 1;
QY 1 attatgtggaaactaccctcgattctctgctgcagagcagcgtcgcgcttccacccc 60
Db 50 ATTATGTGGAATCTACCTCGGATTCCTGCTGCCAGAGCGGCCGCGCTTCCACCGC 109
QY 61 agtcagcgtctccctcgtggtggtgaaagagactcgggagtcgctctccaaagtgc 120
Db 110 AGCGCAGCCTTCCCGCGCTGGGCTGA----GCTTTGGAGTCTGCTTCCCGAGTGC 164
QY 121 cgcgtgagtgagctctcaccacagtcagccaaatgaagcctcttcgggcttctcctg 180
Db 165 CGCGCGGAGTGAGCCTTCGCCAGTCAAGCAATGCTCTCTCGGCCCTCTCTCTG 224
QY 181 acatctgcccctgcccgcagagacagggactcagcagcgaatccaaacctgagtagtaa 240
Db 225 ACATCTGCCCTGCGCGCCAAAGACGCGGACTCGGGCTGAGTCCCACTTGACGACGA 284
QY 241 ttccagttttccagcaacaagggaacgagcggagtagacaagatcctcagcagtagaga 300
Db 285 TTGAGCTCTCCAGCGCAAGGAACAGACGAGTGCAGATCCCGGCATGAGAGATT 344
QY 301 attactgtctactaatgaagattaccacagcccccaagggttctcctaactatccaaga 360
Db 345 GTCATATATCTGTTAATGGAGCATCCACAGCCCCGAAGTTTCTCTATAGTACCCCA 404
QY 361 aatacggctgtgtgagattagtagcagtagagagaaaatgtatgatacaactcacg 420
Db 405 AATATGGTGTGTGTGGAGATTAGTTGCGAGTAGATAAATGTGCGGATCCAGCTGACA 464
QY 421 ttgatgaaagatttgggcttgaagaccaggaagatgacatagcaagtagatatttga 480
Db 465 TTTGATGAGAGATTGGGCTGGAAGATCCAGAACGATATATCAAGTATGATTTTGTGA 524
QY 481 gaagtggagaaacccagtgatggaactatattaggcgctggtggtctgtactgta 540
Db 525 GAAGTTGAGGAGCCCCAGTGATGGAAGCGTTTATAGGACGCTGGTGTGGGACTGTG 584
QY 541 ccaggaaacagatttcttaaggaatcaaattagataagatttgcacacatttcacagaa 600
Db 585 CCAGGAAGCAGACTTCTAAAGGAATCATATCAGGATAAGATTGTATCTGATGAGTAT 644
QY 601 ttctcttgaaacaggggttctgcctccatccacaaattgtcgcacaaatttcacagaa 660
Db 645 TTTCCATCTGAACCCCGATTCTGCATCCACTACAGTATTATCATGCCACAAGTCACAGAA 704
QY 661 gctgtgagctctcagtgctacccctcagcttggccactggacctgcttaataatgct 720
Db 705 ACCACGAGTCTTTCGGTGTGGTCCCTTTCATCTTGTGATTTGGACCTGCTCAACAATG 764
QY 721 ataactgccttagtagaccttggaagaccttcttcgatctcttgaaaccagagagatggcag 780
Db 765 GTGACTTGCCTTCAGTACCTTGGAGAGAGCTGATTGCGGTACCTAGCCAGATCATGCGCAG 824
QY 781 ttggacttagaagatctctatataggccaacttgcaactctctggcaagcttttggttttt 840
Db 825 GTGACTTGGACAGCCTCTACAAAGCAACATGCGACGCTTTTGGCAAGGCTTTCCCTGTAT 884
QY 841 ggaagaaatccagagtggtggtactgaaacctctcaacagagagagtgtaagattatcacgc 900
Db 885 GGGAAAAAAGCAAGGAGTGGTGAATCTCTCTCTCAAGGAGAGGTTAAACCTCTACAGC 944

QY 901 tgcacacctogtaactctcagtggtccatcaagggaagaactaaagaaacogataccatt 960
Db 945 TGCACACCCCGGAACCTCTCAGTGTCCATACGGGAAGAGCTAAAGAGGACAGATACCATA 1004
QY 961 ttctggccaggtgtctcctcgtggttaaacgctgtggtgggaactgtcctgtgtctccac 1020
Db 1005 TTTGCGCCAGTGTCTCTCCGGTCAACGCTGTGGAGAAATTTGCTGTGTCTCCAT 1064
QY 1021 aattgcaatgaatgcaatgtctcccaagcaaatgtactaaaataaccacagagtcctt 1080
Db 1065 AATTGCAATGAATGTAGTGTCTCCACGTAAAGTTACAAAAAGTACCATGAGTCCCT 1124
QY 1081 cagttgagacccaagacggtgtcaggggattgcacaaatacaactcacagcgtgccttg 1140
Db 1125 CAGTTGAGACCAAAAACCTGGAGTCAAGGGATTGCATAGTCACTGATGTGCTCTG 1184
QY 1141 gacacacatgagagtgctgactgtgtgcagagggagcagagagatagccacatcac 1200
Db 1185 GAACACACAGGAATGTGACTGTGTGTGTAGAGGAAACGAGGAGGTAACCTGCAGCCT 1244
QY 1201 caccagcagctc 1212
Db 1245 TCCTAGCAGCAC 1256

RESULT 10
LOCUS AB033830 1116 bp mRNA ROD 10-FEB-2001
DEFINITION Rattus norvegicus rScdglf mRNA for spinal cord-derived growth factor, complete cds.
ACCESSION AB033830
VERSION AB033830.1 GI:11994799
KEYWORDS spinal cord-derived growth factor; SCDF.
SOURCE Rattus norvegicus (strain:Wistar) Adult Kidney cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
1 (sites)
Hamada,T., Ui-Tel,K., Imaki,J. and Miyata,Y.
Molecular Cloning of SCDF-B, a Novel Growth Factor Homologous to SCDF/PDGF-C/Fallotin
Biochem. Biophys. Res. Commun. 280 (3), 733-737 (2001)
11162582
2 (bases 1 to 1116)
Hamada,T., Ui-Tel,K. and Miyata,Y.
Direct Submission
Submitted (25-Oct-1999) Tsuyoshi Hamada, Nippon Medical School, Department of Pharmacology; 1-1-5, Sendagi, Bunkyo-ku, Tokyo, 113-8602, Japan (E-mail:t.hamada@nms.ac.jp, Tel:81-3-3822-2131(ex.5277), Fax:81-3-5814-1684)
Location/Qualifiers
1. .1116
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/tissue_type="Kidney"
/dev_stage="Adult"
1. .1038
/gene="rScdglf"
1. .1038
/gene="rScdglf"
/codon_start=1
/product="spinal cord-derived growth factor"
/protein_id="BAB19969.1"
/db_xref="GI:11994800"
/translation="MLLLGLLLTSLALAGORTGTRAESNLSSKLQSSDKEQNVQDP
RHERVVTSGNSIHSPKPHPTYPNTLVLRVAVDENVRIOITFDERFGLDEDD
LKYQVTEEPESDGLVGCWGTVPKGTSGNHIRIRFVSDYFPPSEDFCFIHY
STIMQVETTPSPVPSALLDLNNAVAFSTVEELIRFLEPDRWQIDLSLVKP
TWPLLGKFLYGGKSKAVNLNLLKEVKLYSCTPRNFSVSIREELKRTDIFWPGCLL
VRCGGNCACCLHNCNEQCQVPRKVTYKTHVQLRPKIGVKGLHLKSLTDVALEHHEE

FEATURES
source

gene

CDS

CDCVCRGNTEG*

BASE COUNT 301 a 267 c 286 g 262 t
ORIGIN

Query Match 44.2%; Score 780; DB 94; Length 1116;
Best Local Similarity 83.5%; Pred. No. 5.9e-198;
Matches 885; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 154 atgagcccttcgggctctcctcgtgacatcgtccctggtccgagcagagagggact 213
Db 1 ATGCTCCCTCGGCTCTCCTCTGCTGACATCTGCCTGGCCGCCAGAGAACC GGAGCT 60
QY 214 caggcggaatccaacctgagtagtaaatccagtttccagcaacaagaagaaacagacga 273
Db 61 CGGGCTGAGTCCACCTGAGCAGCAAGTTGACGCTCTCCAGCGACAAAGAGCAGAACGA 120
QY 274 gtacaagatcctcagcatgagagaattattactgtgtctactaatggaagtattcacagc 333
Db 121 GTGCAAGATCCCGCATGAGAGAGTTGTCACTATATCTGGTAATGGGAGCATCCACAGC 180
QY 334 ccaaggtttcctcaccattatccaagaaatacgggtcttggtatgagagattagtagcagta 393
Db 181 CCGAAGTTTCCTCATACCTTACCTACCCAAAGAAATACGGTCTGTGCTGGAGATTAGTGCAGTA 240
QY 394 gaggaataatgtagatacaaccttacgtttgatgaaagatttggccttgaagaccacgaa 453
Db 241 GATGAAATGTCCGATCCAGCTGACATTTGATGAGAGATTTGGGCTGGAAACCCAGAG 300
QY 454 gatgacatatgcaagtgtgatttttagaagttgaggaacccagtgatggaactatatta 513
Db 301 GACGATCTATGCAAGTATGATTTTGTAGAAGTTGAGGAGCCCGAGTGATGGAAGTGTATTA 360
QY 514 gggcctggtgtgtctggtactgtaccaggaacaaacagatttctaaaggaatacaaat 573
Db 361 GGACGCTGTGTGTCTGGGACTGTGCCAGGAAAGCAGACTTCTAAAGGAAATCATATC 420
QY 574 aggataaattgtatcgtatgaataatttctccttgaaacagggttctgcaccactac 633
Db 421 AGGATAAGATTGTATCTGATGATGATTTCCCATCTGAACCCGGATTTGCAATCCACTAC 480
QY 634 aacattgtcatgccacaattccagaagctgtgagtccttcagtgctaccccttcagct 693
Db 481 AGTATATCATGCCACAAGTACACAGAACTACGAGTCTCTCAGTGTGGCCCTTCAGCT 540
QY 694 ttggcactggacctgttaataatgctataactgctttagtaccttggagacacttatt 753
Db 541 TTGTCATTGGACCTGCTCAATAATGCTGTGACTGCTTTCAGTACTTGTGAAGAGCTGATT 600
QY 754 cgatatcttgaaacagagagatggcagttgagcttagaagatctatataggccaacttgg 813
Db 601 CGGTTCCTAGAGCCAGATCGATGGCAGATAGACTTTGGACAGCTCTACAAAGCAACATGG 660
QY 814 caactcttggcaagccttttttttggaaagaaacccagagtggtgatctgaacctt 873
Db 661 CCGCTTTTGGGCAAGCTTTCTCTGTATGGGAAAAAAGCAAGCGGTGAATCTGAACCTT 720
QY 874 ctacaagagaggtgaagattatcacagctgcacacctcgtaacttctcagtgctccataagg 933
Db 721 CTCAAAGAAGAGGTAAACTTTTACAGCTCCACACCCCGGAACCTTCTCTGTGTCCATACGG 780
QY 934 gaagaactaaagaaacccagataccattttctggcaggtgtgtcctcgtgttaaacgctgt 993
Db 781 GAAGAGCTAAAGAGGACAGACACACCATATTCTGGCCAGGTTGTCTCCTGGTCAAGCGCTGT 840
QY 994 ggtgggaactgtgcctgtgtctccacaattgcagaatgctcaatgtgtcccaagcaaa 1053
Db 841 GGAGGAAATGTGCTGTGTGCTCCATTAATGCAATGAATGTGAGTGTGTCTCCACGTAAG 900
QY 1054 gttactaaaataaccagaggtctctcagttgagacaaacagccggtgtcaggggattg 1113
Db 901 GTTACAAAAAAGTACCATGAGGTCCTCTCAGTTGAGACCAAAAATTTGGAGTCAAGGGATTG 960

VERSION KEYWORDS SOURCE

AC009582.4 GI:7658376
HTG: HTGS_PHASE1; HTGS_DRAFT.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 191314)

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 4, clone RP11-154F14

REFERENCE

2 (bases 1 to 191314)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,

Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A.,

Castle, A., Cerny, J., Collangelo, M., Collins, S., Collymore, A.,

Cooke, P., DeRellano, K., Depayre, E., Devon, K., Dewar, K.,

Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C.,

Funk, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,

Hagos, B., Heaford, A., Horton, L., Howland, J.C., Jones, C., Kann, L.,

Karatas, A., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P.,

Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,

Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,

Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,

Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,

Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,

Tesfaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,

Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.

Direct Submission

Submitted (27-AUG-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 28, 2000 this sequence version replaced gi:7381775.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1726

Center clone name: 154_F_14

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 173370 bases at least Q40

Consensus quality: 182196 bases at least Q30

Consensus quality: 185724 bases at least Q20

Insert size: 194000; agarose-fp

Insert size: 189414; sum-of-contigs

Quality coverage: 3.6 in Q20 bases; agarose-fp

Quality coverage: 3.7 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 30 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 1155: contig of 1155 bp in length

1156 1255: gap of 100 bp

1256 2550: contig of 1295 bp in length

2551 2650: gap of 100 bp

2651 4185: contig of 1535 bp in length

4186 4285: gap of 100 bp

4286 5873: contig of 1588 bp in length

5874 5973: gap of 100 bp

5974 7991: contig of 2018 bp in length

7992 8091: gap of 100 bp

8092 10978: contig of 2887 bp in length

10979 11078: gap of 100 bp

* 11079 13450: contig of 2372 bp in length

* 13451 13550: gap of 100 bp

* 13551 16422: contig of 2872 bp in length

* 16423 16522: gap of 100 bp

* 16523 18672: contig of 2150 bp in length

* 18673 18772: gap of 100 bp

* 18773 22137: contig of 3365 bp in length

* 22138 22237: gap of 100 bp

* 22238 25136: contig of 2899 bp in length

* 25137 25236: gap of 100 bp

* 25237 28475: contig of 3239 bp in length

* 28476 28575: gap of 100 bp

* 28576 31896: contig of 3321 bp in length

* 31897 31996: gap of 100 bp

* 31997 34238: contig of 2242 bp in length

* 34239 34338: gap of 100 bp

* 34339 39364: contig of 5026 bp in length

* 39365 39464: gap of 100 bp

* 39465 44334: contig of 4870 bp in length

* 44335 44434: gap of 100 bp

* 44435 49563: contig of 5129 bp in length

* 49564 49663: gap of 100 bp

* 49664 57336: contig of 7673 bp in length

* 57337 57436: gap of 100 bp

* 57437 63674: contig of 6238 bp in length

* 63675 63774: gap of 100 bp

* 63775 70889: contig of 7115 bp in length

* 70890 70989: gap of 100 bp

* 70990 78191: contig of 7202 bp in length

* 78192 78291: gap of 100 bp

* 78292 85646: contig of 7355 bp in length

* 85647 85746: gap of 100 bp

* 85747 93984: contig of 8238 bp in length

* 93985 94084: gap of 100 bp

* 94085 103005: contig of 8921 bp in length

* 103006 103105: gap of 100 bp

* 103106 114072: contig of 10967 bp in length

* 114073 114172: gap of 100 bp

* 114173 126541: contig of 12369 bp in length

* 126542 126641: gap of 100 bp

* 126642 140995: contig of 14354 bp in length

* 140996 141095: gap of 100 bp

* 141096 154380: contig of 13285 bp in length

* 154381 154480: gap of 100 bp

* 154481 170279: contig of 15799 bp in length

* 170280 170379: gap of 100 bp

* 170380 191314: contig of 20935 bp in length.

FEATURES

source

1. .191314

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="4"

/map="4"

/clone="RP11-154F14"

/clone_lib="RPC1-11 Human Male BAC"

1. .1155

/note="assembly_fragment"

1256. .2550

/note="assembly_fragment"

2651. .4185

/note="assembly_fragment"

4286. .5873

/note="assembly_fragment"

5974. .7991

/note="assembly_fragment"

8092. .10978

/note="assembly_fragment"

11079. .13450

/note="assembly_fragment"

13551. .16422

/note="assembly_fragment"

16523. .18672

/note="assembly_fragment"

misc_feature

| | | | |
|----|-----|---|-----|
| Qy | 520 | tgggtggttctgtagtgaccaggaaacagatttctaaggagaatcaaataggata | 579 |
| Db | 358 | TGGTGTGGTTCTGCTACTGTACCAAGAAAACAGATTTCTAAAGGAATCAAAATTAGGATA | 417 |
| Qy | 580 | agatttgtatctgatgaatatcttcctctgaaaccaagggtctgcactccactacaacatt | 639 |
| Db | 418 | AGATTTGTATCTGATGAATATTTTCCTTCTGAACCAGGGTTCTGCATCCACTACAACATT | 477 |
| Qy | 640 | gtcatgccacaattcaccagaagtctg | 666 |
| Db | 478 | GTCATGCCACAAATTCAGAAGCTGTG | 504 |

Search completed: June 30, 2001, 01:49:42
Job time: 4214 sec

RESULT 15

| | | | | | |
|------------|---|-------------|-------|-------|-------------|
| AX028057 | AX028057 | 504 bp | DNA | PRI | 16-SEP-2000 |
| LOCUS | Sequence 28 from Patent WO0037641. | | | | |
| DEFINITION | AX028057 | | | | |
| ACCESSION | AX028057.1 | GI:10188867 | | | |
| VERSION | | | | | |
| KEYWORDS | human. | | | | |
| SOURCE | Homo sapiens | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| REFERENCE | 1 (bases 1 to 504) | | | | |
| AUTHORS | Gordon, R.D., Dijkmans, J.J., Sprengel, J.J., Yon, J.R., Xu, J., Gosiowska, A., and Dhanaraj, S.N. | | | | |
| TITLE | Vascular endothelial growth factor-x | | | | |
| JOURNAL | Patent: WO 0037641-A 29-JUN-2000; | | | | |
| | GORDON ROBERT DOUGLAS (BE) ; DIJKMANS JOSIENA JOHANNA HUBER (BE) ; JANSSEN PHARMACEUTICA NV (BE) ; SPRENGEL JORG JURGEN (BE) ; YON JEFFREY ROLAND (BE) ; XU JEAN (US) ; GOSIEWSKA ANNA (US) ; DHANARAJ SRIDEVI NAIDU (US) | | | | |
| FEATURES | Location/Qualifiers | | | | |
| source | 1..504 | | | | |
| | /organism="Homo sapiens" | | | | |
| | /db_xref="taxon:9606" | | | | |
| BASE COUNT | 163 a | 92 c | 112 g | 137 t | |
| ORIGIN | | | | | |

| Query Match | 25.3% | Score 447; | DB 91; | Length 504; |
|-----------------------|---|---------------------|-----------|-------------|
| Best Local Similarity | 100.0%; | Pred. No. 8.7e-109; | | |
| Matches 447; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY 220 | gaatccaacotgtagtagtaataattccagatgttttccagcaacaaggaaacgaacggaggtacaa | 279 | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| Db 58 | GAATCCAACCTGAGTAGTAATAATCCAGTGTTCAGCAACAAGGAACACGAGCGAGTACAA | 117 | | |
| | | | | |
| | | | | |
| | | | | |
| QY 280 | gatcctcagcatgagagaattattactgtctgtactactaatggaagtatttcagagcccaagg | 339 | | |
| | | | | |
| | | | | |
| | | | | |
| Db 118 | GATCCTCAGCATGAGAGAAATTTATTACTGTGTCTACTAATGGAAATTTTCACAGCCCCAAGG | 177 | | |
| | | | | |
| | | | | |
| | | | | |
| QY 340 | tttcctcatacttcatccaagaatacacggtcttgtagtgagattagtagcagtagagggaa | 399 | | |
| | | | | |
| | | | | |
| | | | | |
| Db 178 | TTTCCCTCATCTTATCCAGAAATACCGCTCTTGATATGGAGATTATGATGCAGTAGAGGAA | 237 | | |
| | | | | |
| | | | | |
| | | | | |
| QY 400 | aatgtatggatacaacttacggtttgatgaaagatttgggcttgaagaccaccaagaagatgac | 459 | | |
| | | | | |
| | | | | |
| Db 238 | AATGTATGGATACAACCTTACGTTTGATGAAAGATTTGGGCTTGAAGCCCCAGAGAATGAC | 297 | | |
| | | | | |
| | | | | |
| | | | | |
| QY 460 | atatgcaagtatgattttagaagttgaggaaccccgtagtgaaactatatattagggcgc | 519 | | |
| | | | | |
| | | | | |
| | | | | |
| Db 298 | ATATGCAAGTATGATTGTTAGAAAGTTGAGGAACCCAGTATGGAAGTATATTAGGCGC | 357 | | |
| | | | | |
| | | | | |
| | | | | |

7

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2001, 00:12:33 ; Search time 101.97 Seconds
(without alignments)
10862.199 Million cell updates/sec

Title: US-09-457-066-1
Perfect score: 1764
Sequence: 1 attatgtgaaactaccctg.....cttgtctgctgatagga 1764

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues
Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601:*

1: /SIDS8/gcgdata/geneseq/geneseqn/NA1980.DAT:*

2: /SIDS8/gcgdata/geneseq/geneseqn/NA1981.DAT:*

3: /SIDS8/gcgdata/geneseq/geneseqn/NA1982.DAT:*

4: /SIDS8/gcgdata/geneseq/geneseqn/NA1983.DAT:*

5: /SIDS8/gcgdata/geneseq/geneseqn/NA1984.DAT:*

6: /SIDS8/gcgdata/geneseq/geneseqn/NA1985.DAT:*

7: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT:*

8: /SIDS8/gcgdata/geneseq/geneseqn/NA1987.DAT:*

9: /SIDS8/gcgdata/geneseq/geneseqn/NA1988.DAT:*

10: /SIDS8/gcgdata/geneseq/geneseqn/NA1989.DAT:*

11: /SIDS8/gcgdata/geneseq/geneseqn/NA1990.DAT:*

12: /SIDS8/gcgdata/geneseq/geneseqn/NA1991.DAT:*

13: /SIDS8/gcgdata/geneseq/geneseqn/NA1992.DAT:*

14: /SIDS8/gcgdata/geneseq/geneseqn/NA1993.DAT:*

15: /SIDS8/gcgdata/geneseq/geneseqn/NA1994.DAT:*

16: /SIDS8/gcgdata/geneseq/geneseqn/NA1995.DAT:*

17: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT:*

18: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:*

19: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:*

20: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:*

21: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:*

22: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|--------|---------------|--------|----|-----------------------------|
| 1 | 1746 | 99.0 | 1760 | 21 | AAC81582 Human zveg3 DNA, |
| 2 | 1746 | 99.0 | 1760 | 21 | AA51498 Human growth facto |
| 3 | 1744.4 | 98.9 | 2668 | 21 | AA571990 Human VEGF-X CDNA. |
| 4 | 1734 | 98.3 | 2839 | 21 | AA47452 Human TANGO 128 co |
| 5 | 1732.4 | 98.2 | 2779 | 21 | AA52458 CDNA encoding huma |
| 6 | 1732.4 | 98.2 | 2825 | 20 | AA23691 Human VEGF-E DNA. |
| 7 | 1732.4 | 98.2 | 2849 | 20 | AA334296 Human PRO200 nucle |
| 8 | 1732.4 | 98.2 | 2849 | 21 | AA78582 Human PRO200 (UNQ1 |
| 9 | 1732.4 | 98.2 | 2849 | 21 | AA88515 Human PRO200 (VEGF |
| 10 | 1732.4 | 98.2 | 2849 | 21 | AA58579 Human PRO200 prote |
| 11 | 1732.4 | 98.2 | 2849 | 21 | AA77621 Human PRO713 CDNA |

| | | | | | |
|----|--------|------|------|----|-----------------------------|
| 12 | 1732.4 | 98.2 | 2849 | 22 | AAC90564 Human PRO200 cDNA. |
| 13 | 1732.4 | 98.2 | 2849 | 22 | AAC88962 Human PRO200 codin |
| 14 | 1732.4 | 98.2 | 2849 | 22 | AAC97404 Human angiogenesis |
| 15 | 1729.2 | 98.0 | 2858 | 21 | AZ48599 Bone morphogenic p |
| 16 | 1718.2 | 97.4 | 2776 | 21 | AA71952 Human VEGF-X homol |
| 17 | 1717.2 | 97.3 | 2896 | 21 | AAC64426 Human platelet-der |
| 18 | 1717.2 | 97.3 | 2898 | 20 | AX86352 DNA encoding human |
| 19 | 1689.4 | 95.8 | 2475 | 21 | AA71951 Human RACE generat |
| 20 | 1662.2 | 94.2 | 3087 | 21 | AAF18314 Lung cancer associ |
| 21 | 1616.8 | 91.7 | 2794 | 21 | AZ48600 Bone morphogenic p |
| 22 | 1610.8 | 91.3 | 2108 | 21 | AA12523 cDNA encoding plat |
| 23 | 1497 | 84.9 | 1536 | 21 | AA12524 cDNA encoding a fr |
| 24 | 1356 | 76.9 | 1473 | 21 | AA71955 Human VEGF-X DNA i |
| 25 | 1038.4 | 58.9 | 1134 | 21 | AA71983 Human VEGF-X DNA f |
| 26 | 998 | 56.6 | 1095 | 21 | AA51540 MBP-zVEGF3 fusion |
| 27 | 977 | 55.4 | 1096 | 21 | AA71985 Human VEGF-X DNA f |
| 28 | 973 | 55.2 | 1134 | 21 | AA71984 Human VEGF-X DNA f |
| 29 | 883.2 | 50.1 | 3571 | 21 | AAC81583 Mouse zveg3 DNA, |
| 30 | 883.2 | 50.1 | 3571 | 21 | AA51527 Murine vascular en |
| 31 | 878.8 | 49.8 | 1474 | 21 | AA12525 cDNA encoding muri |
| 32 | 871 | 49.4 | 1878 | 21 | AA71936 Human VEGF-X DNA # |
| 33 | 857.4 | 48.6 | 1716 | 21 | AA71982 Human VEGF-X DNA # |
| 34 | 750 | 42.5 | 822 | 21 | AA71981 Human VEGF-X DNA # |
| 35 | 737.2 | 41.8 | 1035 | 21 | AA51499 Human growth facto |
| 36 | 648.2 | 36.7 | 1710 | 21 | AA71989 Human VEGF-X DNA e |
| 37 | 470 | 26.6 | 542 | 21 | AA71988 Human VEGF-X DNA e |
| 38 | 447 | 25.3 | 550 | 21 | AA71987 Human VEGF-X CUB-1 |
| 39 | 373.4 | 21.2 | 764 | 21 | AA47478 Murine TANGO 128 c |
| 40 | 350 | 19.8 | 500 | 21 | AA71986 Human VEGF-X PDGF- |
| 41 | 283.4 | 16.1 | 289 | 21 | AA97809 Human EST clone DN |
| 42 | 283.4 | 16.1 | 289 | 21 | AA97832 Human EST clone DN |
| 43 | 279 | 15.8 | 279 | 21 | AA97807 Human EST clone DN |
| 44 | 279 | 15.8 | 279 | 21 | AA97828 Human EST clone DN |
| 45 | 278.6 | 15.8 | 284 | 21 | AA97802 Human EST clone DN |

ALIGNMENTS

RESULT 1

AAC81582
ID AAC81582 standard; DNA; 1760 BP.

XX AAC81582;

DT 09-MAR-2001 (first entry)

XX Human zveg3 DNA, SEQ ID NO:32.

XX Human; zveg3; zveg4 fusion; growth factor homologue; VEGF/PDGF family;
CUB domain; PDGF-like activity; mitogenic; osteogenic;
neovascularisation; tissue repair; proliferation; differentiation;
liver damage; neurodegenerative; Alzheimer's disease; multiple sclerosis;
periodontal disease; bone fracture; wound healing; ischaemia;
immunomodulation; hepatic; ds.

OS Homo sapiens.

PN WO200066736-A1.

PD 09-NOV-2000.

XX 03-MAY-2000; 2000WO-US40047.

XX 03-MAY-1999; 99US-0304216.

PR 10-NOV-1999; 99US-0164463.

PR 04-FEB-2000; 2000US-0180169.

XX (ZYMO) ZYMOGENETICS INC.

PI Gilbert T, Hart CE, Sheppard PO, Gilbertson DG;

XX WPI; 2000-687541/67.

DR P-PSDB; AAB48657.
XX Growth factor homologs and the nucleic acids that encode them, useful
PT e.g. for treating liver damage, ischemia, multiple sclerosis and
PT Alzheimer's disease -
XX
PS Claim 25; Page 123-125; 143pp; English.
XX
CC The invention relates to the human growth factor homologue zvegfg4
CC (AAB48653), and nucleic acids encoding it (AAC81555). Zvegfg4 is a member
CC of the PDGF (platelet-derived growth factor)/VEGF (vascular endothelial
CC growth factor) family. Zvegfg4 has a growth factor domain (AAB48654)
CC characterised by a PDGF cysteine knot structure, and a CUB domain
CC (AAB48655) which has a beta barrel structure. Zvegfg4 has PDGF-like
CC activity, having mitogenic activity on fibroblasts, vascular smooth
CC muscle cells and pericytes, and has also been shown to stimulate bone
CC growth. The invention also relates to fusion proteins comprising human
CC zvegfg4 or fragments thereof, particularly human zvegfg4/human zvegfg3
CC fusions; expression constructs and host cells comprising human zvegfg4
CC nucleic acids; the recombinant expression of human zvegfg4; an antibody
CC which binds to human zvegfg4 or a fragment thereof; a method of activating
CC a cell-surface PDGF receptor using a zvegfg4-derived polypeptide; a
CC method of modulating the proliferation, differentiation, migration or
CC metabolism of bone cells, comprising exposing bone cells to
CC zvegfg4-derived polypeptides; and a method of detecting a genetic
CC abnormality in the zvegfg4 gene of a patient. Zvegfg4 proteins and derived
CC fragments may be used to stimulate tissue development or repair, or
CC cellular differentiation or proliferation. They are particularly used for
CC the treatment or repair of liver damage, and may also be used to
CC modulate neurite growth (e.g., in the treatment of Alzheimer's disease or
CC multiple sclerosis). Due to their osteogenic activity, they may be used
CC in the treatment of periodontal disease and fractures. They may also be
CC used to enhance expansion and mobilisation of haematopoietic stem cells
CC and endothelial precursor stem cells, which may be useful in the
CC treatment of ischaemia, in wound healing, and in the modulation of the
CC immune system. The present sequence represents DNA encoding human
CC zvegfg3.
XX
SQ Sequence 1760 BP; 494 A; 373 C; 411 G; 482 T; 0 other;

Query Match 99.0%; Score 1746; DB 21; Length 1760;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1760; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Qy 1 attatgtgaaactaccctgcgattctctgctgcaagagactcgagcgctccacc 60
Db 1 attatgtgaaactaccctgcgattctctgctgcaagagactcgagcgctccacc 60
Qy 61 agtcagccttccctggcggtggtgaaagagactcgagcgctccaccagtcc 120
Db 61 agtcagccttccctggcggtggtgaaagagactcgagcgctccaccagtcc 120
Qy 121 cgcctgagtgagctcaccacagtcacgaataagagcctcttccttcctgctg 180
Db 121 cgcctgagtgagctcaccacagtcacgaataagagcctcttccttcctgctg 180
Qy 181 acatctgcctggcgccagagacagggactcagcggaatccaaacctgagtagtaa 240
Db 181 acatctgcctggcgccagagacagggactcagcggaatccaaacctgagtagtaa 240
Qy 241 ttccagtttccagcaacaagaacagagagagactcagagatccctcagcagatgagaatt 300
Db 241 ttccagtttccagcaacaagaacagagagagactcagagatccctcagcagatgagaatt 300
Qy 301 attactgtctactaataagagattatccacagcccaaggttctcctactatccaaaga 360
Db 301 attactgtctactaataagagattatccacagcccaaggttctcctactatccaaaga 360
Qy 361 aatcagctcttggtatgagattagtagcagtagaggaataatgtagtatacaacttacc 420
Db 361 aatcagctcttggtatgagattagtagcagtagaggaataatgtagtatacaacttacc 420

Qy 421 ttgatgaagagatttgggttgaaagaccagaaagatgacatatgcaagtagtatttctga 480
Db 421 ttgatgaagagatttgggttgaaagaccagaaagatgacatatgcaagtagtatttctga 480
Qy 481 gaagttagaaccacagtgatgaactatattagggcgctggtggttcttgactgta 540
Db 481 gaagttagaaccacagtgatgaactatattagggcgctggtggttcttgactgta 540
Qy 541 ccaggaaaaacagatttctaaaggaaatcaaataggatgaagatttctctgataaatat 600
Db 541 ccaggaaaaacagatttctaaaggaaatcaaataggatgaagatttctctgataaatat 600
Qy 601 ttctctctgaaccaggggttctgcatccactacacattgttcacacaaattccacagaa 660
Db 601 ttctctctgaaccaggggttctgcatccactacacattgttcacacaaattccacagaa 660
Qy 661 gctgtgagtcctcagtgctacccttcagcttgcactggacctgcttaataatgct 720
Db 661 gctgtgagtcctcagtgctacccttcagcttgcactggacctgcttaataatgct 720
Qy 721 ataactcctttagtccttggaaagaccttattcgatatcttgaaacagagatggcag 780
Db 721 ataactcctttagtccttggaaagaccttattcgatatcttgaaacagagatggcag 780
Qy 781 ttgacttagaagatctatatagcccaactggcaacttcttgcaagggtttgttttt 840
Db 781 ttgacttagaagatctatatagcccaactggcaacttcttgcaagggtttgttttt 840
Qy 841 ggaagaaaaatccagagtggtggaactgaaccttcaacagagagtaagatatcacgc 900
Db 841 ggaagaaaaatccagagtggtggaactgaaccttcaacagagagtaagatatcacgc 900
Qy 901 tgcaacctcgttaacttctcagtgctcacaaggaagaaactaaagaaacccgataccatt 960
Db 901 tgcaacctcgttaacttctcagtgctcacaaggaagaaactaaagaaacccgataccatt 960
Qy 961 ttctggccaggttctcctcctgggttaaacgctgtgtgggaactgtgctgtgtctccac 1020
Db 961 ttctggccaggttctcctcctgggttaaacgctgtgtgggaactgtgctgtgtctccac 1020
Qy 1021 aattgcaatgaatgcaatgtgtcccaagcaaaagttaactaaataaccagaggtcct 1080
Db 1021 aattgcaatgaatgcaatgtgtcccaagcaaaagttaactaaataaccagaggtcct 1080
Qy 1081 cagttagacccaagaccggtgtcaggggttcacaaactcaactcaccgctggccctg 1140
Db 1081 cagttagacccaagaccggtgtcaggggttcacaaactcaactcaccgctggccctg 1140
Qy 1141 gacacacatgagagtgactgtgtgtagagggagacagagagatagccgcatcac 1200
Db 1141 gacacacatgagagtgactgtgtgtagagggagacagagagatagccgcatcac 1200
Qy 1201 caccagcagcttctccacagctgtgcagtllylvgcagtggtgacttattagagaac 1260
Db 1201 caccagcagcttctccacagctgtgcagtllylvgcagtggtgacttattagagaac 1260
Qy 1261 gtatggttatctccatcccttaactcctcagttgttgcctcaagaccttctcctcagg 1320
Db 1261 gtatggttatctccatcccttaactcctcagttgttgcctcaagaccttctcctcagg 1320
Qy 1321 atttacagtgctattctgaagagagagacatcaacagaatttagagttgtgcaacagctc 1380
Db 1321 atttacagtgctattctgaagagagagacatcaacagaatttagagttgtgcaacagctc 1380
Qy 1381 ttctgagagagagcctaaag 1440
Db 1381 ttctgagagagagcctaaag 1440
Qy 1441 tgtgtattataatagatcacagctagtttcagagttaccatgactatccactagct 1500
Db 1441 tgtgtattataatagatcacagctagtttcagagttaccatgactatccactagct 1500
Qy 1501 ggggtctgtatttctcagttcttcgatacagcgcttagggtaagtacagggagagagagag 1560
Db 1501 ggggtctgtatttctcagttcttcgatacagcgcttagggtaagtacagggagagagagag 1560

[illegible]

| | | |
|----|----------------------------|--|
| | CC | factor domain and a CUB domain (generic sequence motifs are shown in |
| | CC | AA96859 and AA96860). The growth factor domain is characterized by an |
| | CC | arrangement of cysteine residues and beta-strands that is characteristic |
| | CC | of the "cysteine knot" structure of the platelet-derived growth factor |
| | CC | (PDGF) family. The CUB domain shows homology to CUB domains in |
| | CC | neuropilins, human bone morphogenetic protein-1, porcine seminal plasma |
| | CC | protein, bovine acidic seminal fluid protein and Xenopus laevis |
| | CC | tollid-like protein. Structural analysis and homology predict that |
| | CC | ZVEGF3 polypeptides complex with a second polypeptide to form multimeric |
| | CC | proteins. The human zvegf3 gene has been mapped to chromosome 4q28.3. |
| | CC | ZVEGF3 is useful for stimulating the growth of fibroblasts or smooth |
| | CC | muscles cells, for activating cell surface PDGF-alpha receptor and for |
| | CC | inhibiting PDGF-alpha receptor mediated cellular processes. ZVEGF3 is |
| | CC | useful for regulating (post-development) organ growth, regeneration and |
| | CC | maintenance, as well as tissue maintenance and repair processes. ZVEGF3 |
| | CC | antagonists are useful for treating cancer, rheumatoid arthritis, |
| | CC | diabetic retinopathy, ischemic limb disease, peripheral vascular |
| | CC | disease, myocardial ischemia, vascular intimal hyperplasia, |
| | CC | atherosclerosis, wound healing, chronic liver disease and haemangioma |
| | CC | formation. ZVEGF3 can also be used to modulate neurite growth and |
| | CC | development of the nervous system, and for treating neurodegenerative |
| | CC | diseases. |
| | XX | |
| SQ | Sequence | 1760 BP; 494 A; 373 C; 411 G; 482 T; 0 other; |
| | Query Match | 99.0%; Score 1746; DB 21; Length 1760; |
| | Best Local Similarity | 99.8%; Pred. No. 0; |
| | Matches 1760; Conservative | 0; Mismatches 0; Indels 4; Gaps 1; |
| Qy | 1 | attatgtggaactaccctcgattcttctgtgccagacaggctgcggcttcacccc 60 |
| Db | 1 | attatgtggaactaccctcgattcttctgtgccagacaggctgcggcttcacccc 60 |
| Qy | 61 | agtgcagcttccctggcgggtgtgtaagaagactcgggagtcgtctccaagtgc 120 |
| Db | 61 | agtgcagcttccctggcgggtgtgtaagaagactcgggagtcgtctccaagtgc 120 |
| Qy | 121 | cgcgtgagtgcgtctcacccagtcagcaaatgacctcttcgggtctctctctg 180 |
| Db | 121 | cgcgtgagtgcgtctcacccagtcagcaaatgacctcttcgggtctctctctg 180 |
| Qy | 181 | acattgccttgcgcgcagagacaggggactcaggcggaataccaacctgagtagtaa 240 |
| Db | 181 | acattgccttgcgcgcagagacaggggactcaggcggaataccaacctgagtagtaa 240 |
| Qy | 241 | ttcagttttccagacaagaagaacagagagtagtacaaagactcctcagatgagagaatt 300 |
| Db | 241 | ttcagttttccagacaagaagaacagagagtagtacaaagactcctcagatgagagaatt 300 |
| Qy | 301 | attactggtctaataatgggaagtattcacagcccaagggttctctcatcttatccaaga 360 |
| Db | 301 | attactggtctaataatgggaagtattcacagcccaagggttctctcatcttatccaaga 360 |
| Qy | 361 | aatacggcttggtatggattagtagcagtagagagaaaatgatgtgatacaacttacg 420 |
| Db | 361 | aatacggcttggtatggattagtagcagtagagagaaaatgatgtgatacaacttacg 420 |
| Qy | 421 | tttgatgaaagatttggccttgaagaccagagaagatgacatatgcaagtatgtatttcta 480 |
| Db | 421 | tttgatgaaagatttggccttgaagaccagagaagatgacatatgcaagtatgtatttcta 480 |
| Qy | 481 | gaagtgaaggaaccagtgatggaactatatattagggcgcgtgggtgttgtctgtactcta 540 |
| Db | 481 | gaagtgaaggaaccagtgatggaactatatattagggcgcgtgggtgttgtctgtactcta 540 |
| Qy | 541 | ccaggaaaaacagatttctaaaggaaatacaaataggataagatttgtatctgtatgaatat 600 |
| Db | 541 | ccaggaaaaacagatttctaaaggaaatacaaataggataagatttgtatctgtatgaatat 600 |
| Qy | 601 | tttctcttggaaccagggttctgcataccaatacaacattgtcatgccaatattcaagaa 660 |
| Db | 601 | tttctcttggaaccagggttctgcataccaatacaacattgtcatgccaatattcaagaa 660 |

[illegible]

| | | | |
|----|---|---------------------|----|
| FH | Key | Location/Qualifiers | |
| FT | CDS | 288..1325 | |
| FT | | /*tag= a | |
| FT | | /product= TANGO | 38 |
| XX | | | |
| PN | WO200039284-A1. | | |
| XX | | | |
| PD | 06-JUL-2000. | | |
| XX | | | |
| PF | 23-DEC-1999; | 99WO-US31025. | |
| XX | | | |
| PR | 30-DEC-1998; | 98US-0223546. | |
| XX | | | |
| PA | (MILL-) MILLENNIUM PHARM INC. | | |
| XX | | | |
| PI | Holtzman DA; | | |
| XX | | | |
| DR | WPI; 2000-465743/40. | | |
| DR | P-PSDB; AAB01419. | | |
| XX | | | |
| PT | Novel nucleic acid sequences encoding TANGO-128, 140, 197, 212, 213, | | |
| PT | 224 and 239 polypeptides useful for the treatment of asthma, rheumatoid | | |
| PT | arthritis, psoriasis and autoimmune diseases | | |
| XX | | | |
| PS | Claim 1; Fig 1; 209pp; English. | | |
| XX | | | |
| CC | Nucleic acids encoding TANGO polypeptides are useful as modulating | | |
| CC | agents for regulating cellular processes like asthma, graft | | |
| CC | versus-host diseases, rheumatoid arthritis, psoriasis, inflammatory | | |
| CC | bowel disease, septic shock, ulcerative colitis, Crohn's disease, | | |
| CC | chronic myelogenous leukemia, cancer, liver disease, Hodgkin's | | |
| CC | disease, osteoarthritis, Lyme's disease, cachexia and autoimmune | | |
| CC | diseases e.g. myasthenia gravis, autoimmune diabetes and systemic | | |
| CC | lupus erythematosus. The nucleic acids are also useful for producing | | |
| CC | transgenic animals and the TANGO polypeptides themselves. Partial | | |
| CC | TANGO-128, 140, 197, 212, 213, 224, 239 sequences are useful in | | |
| CC | forensic biology, for diagnostic assays, prognostic assays, | | |
| CC | pharmacogenomics and for monitoring clinical trials. TANGO | | |
| CC | polypeptides are suitable for both prophylactic and therapeutic | | |
| CC | methods for treating a subject at risk of a disorder or having a | | |
| CC | disorder associated with aberrant TANGO expression. A wide range | | |
| CC | of cellular disorders can be treated. | | |
| XX | | | |
| SQ | Sequence 2839 BP; 858 A; 532 C; 600 G; 849 T; 0 other; | | |
| | | | |
| | Query Match 98.3%; Score 1734; DB 21; Length 2839; | | |
| | Best Local Similarity 99.7%; Pred. No. 0; | | |
| | Matches 1759; Conservative 0; Mismatches 0; Indels 5; Gaps 2; | | |
| Qy | 1 attatgtgaaactaccctgcgattctctgtgccagagcaggctcgcgcttccacccc | 60 | |
| Db | | | |
| Qy | 135 attatgtgaaactaccctgcgattctctgtgccagagcaggctcgcgcttccacccc | 194 | |
| Qy | 61 agtcagccttccctgcggtggtgaaagagactcgggagtcgctgtccaaagtgc | 120 | |
| Db | | | |
| Qy | 195 agtcagccttccctgcggtggtgaaagagactcgggagtcgctgtccaaagtgc | 254 | |
| Qy | 121 qcctgagtagctctaccctcagcgaatgagcctcttcggcttctctctgtg | 180 | |
| Db | | | |
| Qy | 255 qcctgagtagctctaccctcagcgaatgagcctcttcggcttctctctgtg | 314 | |
| Qy | 181 acatctgcctgcgcccagagacagggactcagggcggaatccaaacctgagtagtaa | 240 | |
| Db | | | |
| Qy | 315 acatctgcctgcgcccagagacagggactcagggcggaatccaaacctgagtagtaa | 374 | |
| Qy | 241 ttccagtttccagcaacaagaacagacggagtagtagatctctcagcatgagagaatt | 300 | |
| Db | | | |
| Qy | 375 ttccagtttccagcaacaagaacagacggagtagtagatctctcagcatgagagaatt | 434 | |
| Qy | 301 attactgtctactaatggaagtattcacagcccaaggtttctcattatccaaaga | 360 | |
| Db | | | |
| Qy | 435 attactgtctactaatggaagtattcacagcccaaggtttctcattatccaaaga | 494 | |
| Qy | 361 aatacggctctgtgtatggagattagtagcagtagagagaaatgtatggtacacacttacg | 420 | |
| Db | | | |
| Qy | 495 aatacggctctgtgtatggagattagtagcagtagagagaaatgtatggtacacacttacg | 554 | |
| Db | | | |
| Qy | 421 ttgatgaaagatttggcttgaagaccagagatgacatatgcaagtattgatttga | 480 | |
| Db | | | |
| Qy | 555 ttgatgaaagatttggcttgaagaccagagatgacatatgcaagtattgatttga | 614 | |
| Db | | | |
| Qy | 481 gaagttaggaacccagtgatggaactatatattagggcgctggtgtctctggtactgta | 540 | |
| Db | | | |
| Qy | 615 gaagttaggaacccagtgatggaactatatattagggcgctggtgtctctggtactgta | 674 | |
| Db | | | |
| Qy | 541 ccaggaaaacagatttctaaaggaaaatcaaataggataagatttgtatctgatgaatat | 600 | |
| Db | | | |
| Qy | 675 ccaggaaaacagatttctaaaggaaaatcaaataggataagatttgtatctgatgaatat | 734 | |
| Db | | | |
| Qy | 601 ttctcttgaaacagggttctgcatccactacaacattgtcatgcccacaaatccacagaa | 660 | |
| Db | | | |
| Qy | 735 ttctcttgaaacagggttctgcatccactacaacattgtcatgcccacaaatccacagaa | 794 | |
| Db | | | |
| Qy | 661 gctgtaagctctcagtgctacccttccagcttgcactgacctgacctgcttaataatgct | 720 | |
| Db | | | |
| Qy | 795 gctgtaagctctcagtgctacccttccagcttgcactgacctgacctgcttaataatgct | 854 | |
| Db | | | |
| Qy | 721 ataactgccttttagtaoacttggagaocttattcgatatcttgaaccagagagatggcag | 780 | |
| Db | | | |
| Qy | 855 ataactgccttttagtaoacttggagaocttattcgatatcttgaaccagagagatggcag | 914 | |
| Db | | | |
| Qy | 781 ttggactagaagatctatatagccaaacttgccaaacttcttggaagggtttgttttt | 840 | |
| Db | | | |
| Qy | 915 ttggactagaagatctatatagccaaacttgccaaacttcttggaagggtttgttttt | 974 | |
| Db | | | |
| Qy | 841 ggaagaaaatccagagtggtgagatctgaaccttctacagagaggttaagatatacagc | 900 | |
| Db | | | |
| Qy | 975 ggaagaaaatccagagtggtgagatctgaaccttctacagagaggttaagatatacagc | 1034 | |
| Db | | | |
| Qy | 901 tgacacactcgttaacttctcagtgctcacaaggaagaaactaaagaaacccgataccatt | 960 | |
| Db | | | |
| Qy | 1035 tgacacactcgttaacttctcagtgctcacaaggaagaaactaaagaaacccgataccatt | 1094 | |
| Db | | | |
| Qy | 961 ttctggccagggttctctcctggttaaacgctgtgtgtgggaactgtgctgttctccac | 1020 | |
| Db | | | |
| Qy | 1095 ttctggccagggttctctcctggttaaacgctgtgtgtgggaactgtgctgttctccac | 1154 | |
| Db | | | |
| Qy | 1021 aattgcaatgaatgcaatgtgtcccaagcaaaagtactaaaaataaccacgaggtcctt | 1080 | |
| Db | | | |
| Qy | 1155 aattgcaatgaatgcaatgtgtcccaagcaaaagtactaaaaataaccacgaggtcctt | 1214 | |
| Db | | | |
| Qy | 1081 cagttgagaccaaagaccggtgtcaggggattgcacaaatcactcacccgctggccctg | 1140 | |
| Db | | | |
| Qy | 1215 cagttgagaccaaagaccggtgtcaggggattgcacaaatcactcacccgctggccctg | 1274 | |
| Db | | | |
| Qy | 1141 gacacacatgagagagtgactgtgtgcagaggagacacagagagatagcgcacac | 1200 | |
| Db | | | |
| Qy | 1275 gacacacatgagagagtgactgtgtgcagaggagacacagagagatagcgcacac | 1334 | |
| Db | | | |
| Qy | 1201 caccagcagctctgccacagctgtgcagtllyllygcagtggtcttctattagagaac | 1260 | |
| Db | | | |
| Qy | 1335 caccagcagctctgccacagctgtgtcagtllyllygcagtggtcttctattagagaac | 1390 | |
| Db | | | |
| Qy | 1261 gtagcgttatctccatccttaattctcagttgtttgttccaaggaccttctcttcagg | 1320 | |
| Db | | | |
| Qy | 1391 gtagcgttatctccatccttaattctcagttgtttgttccaaggaccttctcttcagg | 1450 | |
| Db | | | |
| Qy | 1321 attacagtgcatcttgaaagagagagacatcaaacagaattagaggtgtgcaacagctc | 1380 | |
| Db | | | |
| Qy | 1451 attacagtgcatcttgaaagagagagacatcaaacagaattagaggtgtgcaacagctc | 1510 | |
| Db | | | |
| Qy | 1381 ttctgagagagccttaagagacagagaaaagggttctcattcgtggaagaaaataaa | 1440 | |
| Db | | | |
| Qy | 1511 ttctgagagagccttaagagacagagaaaagggttctcattcgtggaagaaaataaa | 1570 | |
| Db | | | |

Db 765 gctgtgagctctcagctcagctcagcttgcactggagctgcttaaatgct 824
QY 721 ataaactgcttagtacccttgaagaccttattcgatatcttgaaccagagatggcag 780
Db 825 ataaactgcttagtacccttgaagaccttattcgatatcttgaaccagagatggcag 884
QY 781 ttgacttagaagatctatataggccaacttggcaacttcttggcaaggctttgtttt 840
Db 885 ttgacttagaagatctatataggccaacttggcaacttcttggcaaggctttgtttt 944
QY 841 ggaagaaatccagagtggtgagctgaaccttcttaacagagagaggttaagattatacagc 900
Db 945 ggaagaaatccagagtggtgagctgaaccttcttaacagagagaggttaagattatacagc 1004
QY 901 tgcacacctgaacttctcagtgccataagggaagaaactaaagagaaacagcattaccatt 960
Db 1005 tgcacacctgaacttctcagtgccataagggaagaaactaaagagaaacagcattaccatt 1064
QY 961 ttctggccaggtgtctctctgttaaacgctgtggtgggaactgtgcctgttgcctccac 1020
Db 1065 ttctggccaggtgtctctctgttaaacgctgtggtgggaactgtgcctgttgcctccac 1124
QY 1021 aattgcaatgaatgtcattgtgtcccaagcaaaagttaataaaataccacagagtcctt 1080
Db 1125 aattgcaatgaatgtcattgtgtcccaagcaaaagttaataaaataccacagagtcctt 1184
QY 1081 cagttgagacaaagacggtgtcaggggttgacaaatcactcaccagacgtgcccctg 1140
Db 1185 cagttgagacaaagacggtgtcaggggttgacaaatcactcaccagacgtgcccctg 1244
QY 1141 gagcaccatgaggtgtgactgtgtgcagagggagacacagagagatagccgcattcac 1200
Db 1245 gagcaccatgaggtgtgactgtgtgcagagggagacacagagagatagccgcattcac 1304
QY 1201 caccagacgtcttgcacagagcgtgagtglygcagtggtggtgattctattagagaac 1260
Db 1305 caccagacgtcttgcacagagcgtgagtglygcagtggtggtgattctattagagaac 1360
QY 1261 gtagctgtattctccatcttaactcagttgttgccttcaaggacccttctcagcag 1320
Db 1361 gtagctgtattctccatcttaactcagttgttgccttcaaggacccttctcagcag 1420
QY 1321 attacagtgcatctctgaagagagagacatcaaacagaaatagagtggtgcaacagctc 1380
Db 1421 attacagtgcatctctgaagagagagacatcaaacagaaatagagtggtgcaacagctc 1480
QY 1381 ttttgagagggcctaaagacaggggaaagtgcttcaatcgtyggaagaaataaa 1440
Db 1481 ttttgagagggcctaaagacaggggaaagtgcttcaatcgtyggaagaaataaa 1540
QY 1441 tgttgattaaatagatcaccagtagtttcagagttaccatgtacgtattccactagct 1500
Db 1541 tgttgattaaatagatcaccagtagtttcagagttaccatgtacgtattccactagct 1600
QY 1501 ggggtctgtatttcagttcttcgatacggcttagggtaagtgcagtcagtcaggaagaaac 1560
Db 1601 ggggtctgtatttcagttcttcgatacggcttagggtaagtgcagtcagtcaggaagaaac 1660
QY 1561 tgtgcaagtgagcaccctgattccgttgccttgccttaactcctaaagcctcctctctgggc 1620
Db 1661 tgtgcaagtgagcaccctgattccgttgccttgccttaactcctaaagcctcctctctgggc 1720
QY 1621 ctgaaatcgataaaatcgatt 1680
Db 1721 ctgaaatcgataaaatcgga-ttttttttttttttttttttttttttttttttttttttt 1779
QY 1681 accgaacattctgtactacaacacctggttttaaaaggaactatgttgcctatgaat 1740
Db 1780 accgaacattctgtactacaacacctggttttaaaaggaactatgttgcctatgaat 1839
QY 1741 taaactgtgtcgtgatagga 1764
|||||

Db 1840 taaactgtgtcgtgatagga 1863

RESULT 6

AAZ23691
ID AAZ23691 standard; DNA; 2825 BP.

XX AC AAZ23691;

XX DT 11-JAN-2000 (first entry)

XX DE Human VEGF-E DNA.

XX KW VEGF-E; human; vascular endothelial cell growth factor; wound repair; treatment; cardiovascular disorder; endothelial disorder; therapy;
KW tissue generation; regeneration; cardiac hypertrophy; cancer; detection;
KW angiogenic disorder; age-related macular degeneration; vascular disease;
KW neovascularization; tumor; gene mapping; ss.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers
CDS 259..1296

XX FT /*tag= a
FT /product= "VEGF-E"

XX PN W09947677-A2.

XX PD 23-SEP-1999.

XX PF 10-MAR-1999; 99WO-US05190.

XX PR 17-MAR-1998; 98US-0040220.

XX PR 02-NOV-1998; 98US-0184216.

XX PA (GETH) GENENTECH INC.

XX PI Ferrara N, Kuo SS;

XX WPI: 1999-580306/49.

XX P-PSDB; AAY33679.

XX PT New growth factor polypeptide useful for treating cardiovascular or
endothelial disorders, e.g. cardiac hypertrophy

XX PS Claim 2; Fig 1; 122pp; English.

XX CC This invention describes the isolation of a novel human vascular
endothelial cell growth factor-E (VEGF-E) polypeptide which has
CC tranquilizer, vulnery and cardiant activity. VEGF-E can be administered
CC therapeutically, especially by expressing encoding polynucleotides, to
CC treat cardiovascular or endothelial disorders in mammals, especially
CC humans. It is useful in wound repair and tissue generation and
CC regeneration, and may especially be used to treat cardiac hypertrophy
CC It can be combined with a carrier in pharmaceutical compositions, which
CC can be administered to treat disorders as above. VEGF-E can be used to
CC screen for antagonists and agonists, and the antagonists administered to
CC treat angiogenic disorders in mammals (especially humans) e.g. cancer or
CC age-related macular degeneration. It can be used to generate antibodies,
CC useful therapeutically as antagonists, as above. The antibodies are also
CC useful to detect VEGF-E polypeptide, especially to diagnose
CC cardiovascular, endothelial or angiogenic disorders in mammals (e.g.
CC by contacting the antibody with a tissue sample and detecting formation
CC of an antibody-VEGF-E polypeptide complex. Polynucleotides encoding
CC VEGF-E can be used to diagnose cardiovascular and endothelial disorders
CC in mammals, by detecting abnormally high or low VEGF-E gene expression in
CC tissue samples. They can also be used to diagnose a disease or
CC susceptibility to a disease related to a mutated form of VEGF-E (e.g. a
CC cardiovascular, endothelial or angiogenic disorder such as a tumor), by
CC detecting a mutation in the VEGF-E-encoding sequence isolated from a
CC sample. They may also be used to produce probes useful to detect related
CC sequences or for gene mapping. This sequence encodes the human VEGF-E


```
PN WO9946281-A2.
XX
PD 16-SEP-1999.
XX
PF 08-MAR-1999; 99WO-US05028.
XX
PR 10-MAR-1998; 98US-0077450.
PR 11-MAR-1998; 98US-0077632.
PR 11-MAR-1998; 98US-0077641.
PR 11-MAR-1998; 98US-0077649.
PR 12-MAR-1998; 98US-0077791.
PR 13-MAR-1998; 98US-0078004.
PR 17-MAR-1998; 98US-0040220.
PR 20-MAR-1998; 98US-0078886.
PR 20-MAR-1998; 98US-0078910.
PR 20-MAR-1998; 98US-0078936.
PR 20-MAR-1998; 98US-0078939.
PR 25-MAR-1998; 98US-0079294.
PR 26-MAR-1998; 98US-0079656.
PR 27-MAR-1998; 98US-0079663.
PR 27-MAR-1998; 98US-0079664.
PR 27-MAR-1998; 98US-0079689.
PR 27-MAR-1998; 98US-0079728.
PR 27-MAR-1998; 98US-0079786.
PR 30-MAR-1998; 98US-0079920.
PR 30-MAR-1998; 98US-0079923.
PR 31-MAR-1998; 98US-0080105.
PR 31-MAR-1998; 98US-0080107.
PR 31-MAR-1998; 98US-0080165.
PR 31-MAR-1998; 98US-0080194.
PR 01-APR-1998; 98US-0080327.
PR 01-APR-1998; 98US-0080328.
PR 01-APR-1998; 98US-0080333.
PR 01-APR-1998; 98US-0080334.
PR 08-APR-1998; 98US-0081049.
PR 08-APR-1998; 98US-0081070.
PR 08-APR-1998; 98US-0081071.
PR 09-APR-1998; 98US-0081195.
PR 09-APR-1998; 98US-0081203.
PR 09-APR-1998; 98US-0081229.
PR 15-APR-1998; 98US-0081817.
PR 15-APR-1998; 98US-0081838.
PR 15-APR-1998; 98US-0081952.
PR 15-APR-1998; 98US-0081955.
PR 21-APR-1998; 98US-0082568.
PR 21-APR-1998; 98US-0082569.
PR 22-APR-1998; 98US-0082700.
PR 22-APR-1998; 98US-0082704.
PR 22-APR-1998; 98US-0082804.
PR 23-APR-1998; 98US-0082767.
PR 23-APR-1998; 98US-0082796.
PR 27-APR-1998; 98US-0083336.
PR 28-APR-1998; 98US-0083322.
PR 29-APR-1998; 98US-0083392.
PR 29-APR-1998; 98US-0083495.
PR 29-APR-1998; 98US-0083496.
PR 29-APR-1998; 98US-0083499.
PR 29-APR-1998; 98US-0083500.
PR 29-APR-1998; 98US-0083545.
PR 29-APR-1998; 98US-0083554.
PR 29-APR-1998; 98US-0083558.
PR 29-APR-1998; 98US-0083559.
PR 30-APR-1998; 98US-0083742.
PR 05-MAY-1998; 98US-0084366.
PR 06-MAY-1998; 98US-0084414.
PR 06-MAY-1998; 98US-0084441.
PR 07-MAY-1998; 98US-0084598.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084627.
PR 07-MAY-1998; 98US-0084637.
PR 07-MAY-1998; 98US-0084639.
PR 07-MAY-1998; 98US-0084640.
PR 07-MAY-1998; 98US-0084643.

PR 13-MAY-1998; 98US-0085323.
PR 13-MAY-1998; 98US-0085338.
PR 13-MAY-1998; 98US-0085339.
PR 15-MAY-1998; 98US-0085573.
PR 15-MAY-1998; 98US-0085579.
PR 15-MAY-1998; 98US-0085580.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085689.
PR 15-MAY-1998; 98US-0085697.
PR 15-MAY-1998; 98US-0085700.
PR 15-MAY-1998; 98US-0085704.
PR 18-MAY-1998; 98US-0086023.
PR 22-MAY-1998; 98US-0086392.
PR 22-MAY-1998; 98US-0086414.
PR 22-MAY-1998; 98US-0086430.
PR 22-MAY-1998; 98US-0086486.
PR 28-MAY-1998; 98US-0087098.
PR 28-MAY-1998; 98US-0087106.
PR 28-MAY-1998; 98US-0087208.
PR 30-JUL-1998; 98US-0094651.
PR 11-SEP-1998; 98US-0100038.
XX
XX (GETH ) GENENTECH INC.
XX
PA Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX
PI WPI; 1999-551358/46.
XX
DR P-PSDB; AAY41766.
XX
DR
XX
PT New secreted and transmembrane polypeptides and their polynucleotides,
PT useful for treating blood coagulation disorders, cancers and cellular
PT adhesion disorders -
XX
XX Claim 2; Fig 206; 530pp; English.
XX
CC The present invention describes secreted and transmembrane polypeptides
CC and their polynucleotides. The nucleotide sequences are useful as
CC sources of probes, primers, for chromosome mapping, and for generation
CC of antisense sequences. They can also be used to create transgenic
CC animals. The proteins can be used to treat a variety of diseases and
CC disorders, depending on their function. Diseases that may be treated
CC include blood coagulation disorders, cancers and cellular adhesion
CC disorders. They may also be used to raise antibodies. AAZ33891 to
CC AAZ34338, and AAY41685 to AAY41774 represent polynucleotide and
CC polypeptide sequence given in the exemplification of the present
CC invention.
XX
SQ Sequence 2849 BP; 851 A; 528 C; 619 G; 850 T; 1 other;

Query Match 98.2%; Score 1732.4; DB 20; Length 2849;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

QY 1 attatgtggaaactaccctgcgattctctgtccagagcaggtcgcgttccacccc 60
Db 132 attatgtggaaactaccctgcgattctctgtccagagcaggtcgcgttccacccc 191
QY 61 agtcagccttccctggcgttggtgaaagagactcggagtcgctgttccaaagtgc 120
Db 192 agtcagccttccctggcgttggtgaaagagactcggagtcgctgttccaaagtgc 251
QY 121 cgccgtgagtgagctctcaccctcagccaaatgagctcttcgggttctcctgctg 180
Db 252 cgccgtgagtgagctctcaccctcagccaaatgagctcttcgggttctcctgctg 311
QY 181 acatctgccttgcggccagagacagggagactcagcgggaatccaacctgagtagtaa 240
Db 312 acatctgccttgcggccagagacagggagactcagcgggaatccaacctgagtagtaa 371
QY 241 ttccagttttccagcaacaaggaacagacgagtagtacaagatcctcagcatgagagaatt 300
Db 372 ttccagttttccagcaacaaggaacagacgagtagtacaagatcctcagcatgagagaatt 431
```

[illegible]

PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Kljavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;
XX
DR WPI: 2000-611443/58.
DR P-PSDB: AAB44322.
XX

XX Novel PRO polypeptides and polynucleotides used in detection methods,
PT to target bioactive molecules to specific cells, and to modulate
PT cellular activities -
XX

XX Claim 2; Fig 206; 636pp; English.

XX AAC78458 to AAC78599 represent polynucleotide and EST (expressed
CC sequence tag) sequences which encode secreted or transmembrane PRO
CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic
CC activity. The polynucleotides and polypeptides can be used for detecting
CC the presence of PRO polypeptides in samples, for linking bioactive
CC molecules to cells and for modulating biological activities of cells,
CC using the polypeptides for specific targeting. The polypeptide targeting
CC can be used to kill the target cells, e.g. for the treatment of cancers.
CC The polypeptide pairs provide specific targeting of bioactive molecules
CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
CC the isolation of the PRO polynucleotide sequences.

XX Sequence 2849 BP; 851 A; 528 C; 619 G; 850 T; 1 other;

Query Match 98.28; Score 1732.4; DB 21; Length 2849;
Best Local Similarity 99.78; Pred. No. 0;
Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

QY 1 attatgtggaactaccctcgattctctgctgcagagcaggtcgcgcttcacccc 60
DB 132 attatgtggaactaccctcgattctctgctgcagagcaggtcgcgcttcacccc 191
QY 61 agtgcagcttccctggcgtggtgaaagagactcggagctcgtcttcacaaagtcc 120
DB 192 agtgcagcttccctggcgtggtgaaagagactcggagctcgtcttcacaaagtcc 251
QY 121 cgcgtgagtgagctctcaccctcagtcagcaaatgagctctcggcttctcctgctg 180
DB 252 cgcgtgagtgagctctcaccctcagtcagcaaatgagctctcggcttctcctgctg 311
QY 181 acatctgacctgcccgcagagacaggggactcagcgggaatcacaactcagtagtaaa 240
DB 312 acatctgacctgcccgcagagacaggggactcagcgggaatcacaactcagtagtaaa 371
QY 241 ttccagttttccagcaacaaggaaacagagagtagtaacaagatcctcagcatgagagaatt 300
DB 372 ttccagttttccagcaacaaggaaacagagagtagtaacaagatcctcagcatgagagaatt 431
QY 301 attactggtctactaatgaaagtattcagagcccaagggttctcactatcaccaga 360
DB 432 attactggtctactaatgaaagtattcagagcccaagggttctcactatcaccaga 491
QY 361 aatcacggtctgtatgagatagtagcagtagcagagaaataatgtagtaacaacttacg 420
DB 492 aatcacggtctgtatgagatagtagcagtagcagagaaataatgtagtaacaacttacg 551
QY 421 ttgtgaaagatttgggcttgaagaccagcaagatgacatgacagtagtatttga 480
DB 552 ttgtgaaagatttgggcttgaagaccagcaagatgacatgacagtagtatttga 611
QY 481 gaagttagagaaaccagtgatgaaactatatattagggcgctgggttctcgttactgta 540
DB 612 gaagttagagaaaccagtgatgaaactatatattagggcgctgggttctcgttactgta 671
QY 541 ccagggaaacagatttctaagaaatacaattaggaataagatttctcgtatgaatat 600
DB 672 ccagggaaacagatttctaagaaatacaattaggaataagatttctcgtatgaatat 731
QY 601 ttctcttgaaccaggggttctgcatccactacacattgtcatgccacaattcacagaa 660

DB 732 ttctcttgaaccaggggttctgcatccactacaacattgtctgcacaaattcacagaa 791
QY 661 gctgtgagtccttcagtgctaccccttcagctttgctccactgagctgcttaataatgct 720
DB 792 gctgtgagtccttcagtgctaccccttcagctttgctccactgagctgcttaataatgct 851
QY 721 ataatgcctttagttacacttggaaagaccttattcgatatcttgaaccagagatggcag 780
DB 852 ataatgcctttagttacacttggaaagaccttattcgatatcttgaaccagagatggcag 911
QY 781 ttggacttagaagatctatataggccaacttggcaacttcttggcaaggctttgttttt 840
DB 912 ttggacttagaagatctatataggccaacttggcaacttcttggcaaggctttgttttt 971
QY 841 ggaagaaataccagagtggtgatctgaaccttctaacagagaggtagattatacagc 900
DB 972 ggaagaaataccagagtggtgatctgaaccttctaacagagaggtagattatacagc 1031
QY 901 tgcaacacctgtaacttctcagtgctcataaggaagaaactaaagaaacccagataccatt 960
DB 1032 tgcaacacctgtaacttctcagtgctcataaggaagaaactaaagaaacccagataccatt 1091
QY 961 ttctggccaggtgtgtctcctggttaaacgctgtggtggaaactgtcctgtgtgtctccac 1020
DB 1092 ttctggccaggtgtgtctcctggttaaacgctgtggtggaaactgtcctgtgtgtctccac 1151
QY 1021 aattgcaatgaatgcaatgtgtcccaagcaaaagtactaaaaataccacagaggtccctt 1080
DB 1152 aattgcaatgaatgcaatgtgtcccaagcaaaagtactaaaaataccacagaggtccctt 1211
QY 1081 cagtgagaccaaagaccggtgtcaggggatgcaaaaactcactccacgagctggccctg 1140
DB 1212 cagtgagaccaaagaccggtgtcaggggatgcaaaaactcactccacgagctggccctg 1271
QY 1141 gagcaccatgaggtgtgactgtgtgtcagagggagcacagggaggtatagccgcatcac 1200
DB 1272 gagcaccatgaggtgtgactgtgtgtcagagggagcacagggaggtatagccgcatcac 1331
QY 1201 caccagcagcttctgcccagagctgtgcaagtllylvcagtggtctattattagagaac 1260
DB 1332 caccagcagcttctgcccagagctgtgcaagtllylvcagtggtctattattagagaac 1387
QY 1261 gtagcgtttatctcctccttaactcagttgtgttcttcaaggaccttctcatttcagg 1320
DB 1388 gtagcgtttatctcctccttaactcagttgtgttcttcaaggaccttctcatttcagg 1447
QY 1321 atttaccagtgcatcttgaaagaggagacatcaaacagaattaggagttgtgcaacagctc 1380
DB 1448 atttaccagtgcatcttgaaagaggagacatcaaacagaattaggagttgtgcaacagctc 1507
QY 1381 tttttgagaggagagcctaaaggacagagagaaagggttctcaatcgtggaagaaataaa 1440
DB 1508 tttttgagaggagagcctaaaggacagagagaaagggttctcaatcgtggaagaaataaa 1567
QY 1441 tgtgtattataatagatcacacagctagtttcagagttaccatgtacgtattccactagct 1500
DB 1568 tgtgtattataatagatcacacagctagtttcagagttaccatgtacgtattccactagct 1627
QY 1501 ggggttctgtatttccagttcttctcagatcagggcttaggtaattgtcagtagcaagaaaaaac 1560
DB 1628 ggggttctgtatttccagttcttctcagatcagggcttaggtaattgtcagtagcaagaaaaaac 1687
QY 1561 tftgcaagtgaacacactgattccggttgccttgaactcctaaagctccatgtcctctgggc 1620
DB 1688 tftgcaagtgaacacactgattccggttgccttgaactcctaaagctccatgtcctctgggc 1747
QY 1621 ctaaaaatcgataaaaaatctggatttttttttttttttttttttttttttttttttttttt 1680
DB 1748 ctaaaaatcgataaaaaatctggga-ttttttttttttttttttttttttttttttttttttttt 1806
QY 1681 accagaaacattctatgtactacaacacctgggtttttaaaggaactatgttctgtatgaat 1740

Db 1807 accagaaattctatgtactacaaacctggtttttaaaaggaactatgttgctatgaat 1866

QY 1741 taaactgtgtcgtgctgatagga 1764
|||||

Db 1867 taaactgtgtcgtgctgatagga 1890
|||||

RESULT 9
AAA88515
ID AAA88515 standard; cDNA; 2849 BP.
XX
AC AAA88515;
XX
XX 22-JAN-2001 (first entry)
XX
XX Human PRO200 (VEGF-E) cDNA clone DNA29101-1272.
XX
KW PRO200; vascular epithelial growth factor E; VEGF-E; human;
KW ocular disease; retinopathy; maculopathy; therapy;
KW retinitis pigmentosa; macular degeneration; retinal detachment;
KW retinal tear; macular hole; myopia; traumatic choriorretinopathy;
KW acute retinal necrosis syndrome; contusion; edema;
KW retinal vision occlusion; vascular disease; retinal vasculitis;
KW thrombocytopenic purpura; uveitis; retinal occlusion; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 285..1322
FT /*tag= a
FT sig_peptide 285..326
FT /*tag= b
FT mat_peptide 327..1319
FT /*tag= c
XX
XX WO200053760-A2.
XX
XX 14-SEP-2000.
XX
XX 10-MAR-2000; 2000WO-US06319.
XX
XX 12-MAR-1999; 9905-0123957.
XX
XX (GETH) GENENTECH INC.
XX
XX Ferrara N, Goddard A, Gurney AL, Hebert C, Henzel WJ, Kabakoff RC;
PI Klein RD, Kijavini IJ, Kuo SS, La Fleur M, Wood WI;
XX
XX WPI; 2000-587437/55.
XX
XX P-PSDB; AAB19578.
XX
XX Novel PRO polypeptides useful for preventing or rescuing retinal cells
XX from injury caused by ocular diseases such as retinitis pigmentosa,
XX retinopathy, retinal degenerative diseases, degenerative myopia,
XX uveitis -
XX
XX Example 8; Fig 1; 140pp; English.
XX
XX The present sequence is that of cDNA clone DNA29101-1272 or UNQ174
XX (ATCC 209653) isolated from a human glioma cell line G61 cDNA
XX library using probes (see AAA88523-26) based on an expressed sequence
XX tag (see AAA88522) identified on the basis of homology to vascular
XX endothelial growth factor (VEGF). The clone contains a single open
XX reading frame encoding a 345-amino acid protein (see AAB19578)
XX designated PRO200 or VEGF-E. The isolated cDNA can be used in the
XX recombinant production of PRO200 (VEGF-E). The invention relates
XX to the use of PRO polypeptides, including PRO200, to delay, prevent
XX or rescue retinal cells such as retinal neurons selected from
XX photoreceptors, retinal ganglion cells, displaced retinal ganglion
XX cells, amacrine cells, displaced amacrine cells, horizontal and
XX bipolar neurons, and supportive cells (including Mueller cells and
XX pigment epithelial cells) from injury and degradation. The retinal
XX cells are preferably photoreceptors and photoreceptor cell injury or

CC death is caused by retinal injury, light or environmental trauma or
CC by an ocular disease selected from retinitis pigmentosa, macular
CC degeneration, including age-related, retinal detachment, retinal
CC tears, retinopathy, retinal degenerative diseases, macular holes,
CC degenerative myopia, acute retinal necrosis syndrome, traumatic
CC choriorretinopathies or contusion such as Purtscher's retinopathy,
CC edema, ischemic conditions such as central or branch retinal vision
CC occlusion, collagen vascular diseases, thrombocytopenic purpura,
CC uveitis, retinal vasculitis and occlusion associated with Eales
CC disease and systemic lupus erythematosus (claimed).
XX
XX Sequence 2849 BP; 851 A; 528 C; 619 G; 850 T; 1 other;
SQ

Query Match 98.2%; Score 1732.4; DB 21; Length 2849;
Best Local Similarity 99.7%; Pred No. 0;
Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

QY 1 attatgtgaaactaccctgcgattctctgtcgcagagcagctcgcgttcacccc 60
Db 132 attatgtgaaactaccctgcgattctctgtcgcagagcagctcgcgttcacccc 191
QY 61 agtcagccttcctcctgcggtggtgaaagacactcgcgttcacaaagtgc 120
Db 192 agtcagccttcctcctgcggtggtgaaagacactcgcgttcacaaagtgc 251
QY 121 cgcctgagtgagctctcacccagtcagcaaatgagccttcggttcctcgtcg 180
Db 252 cgcctgagtgagctctcacccagtcagcaaatgagccttcggttcctcgtcg 311
QY 181 acatctgcctcgcgcgcagagacagggagctcagcggaatccaaactgagtagtaa 240
Db 312 acatctgcctcgcgcgcagagacagggagctcagcggaatccaaactgagtagtaa 371
QY 241 ttccagtttccagcaacaaggaacagacgagtagtacagatcctcagcatgagagaatt 300
Db 372 ttccagtttccagcaacaaggaacagacgagtagtacagatcctcagcatgagagaatt 431
QY 301 attactgtgtctactaataatggaagtatttcacagcccaagggttccctcactatccaaga 360
Db 432 attactgtgtctactaataatggaagtatttcacagcccaagggttccctcactatccaaga 491
QY 361 aatacggctcttggtatggagattagtagcagtagaggaataatgtaggatacaacttcag 420
Db 492 aatacggctcttggtatggagattagtagcagtagaggaataatgtaggatacaacttcag 551
QY 421 ttgtatgaagaatttggccttgaagaccagagatgacatagcaagtagattttgtga 480
Db 552 ttgtatgaagaatttggccttgaagaccagagatgacatagcaagtagattttgtga 611
QY 481 gaagttgaggaaccagtgatgaaactatattaggcgcgtggtggttctggtactgta 540
Db 612 gaagttgaggaaccagtgatgaaactatattaggcgcgtggtggttctggtactgta 671
QY 541 ccaggaaaaacagatttctaaaggaatacaaataggataagattgtgtatctgataat 600
Db 672 ccaggaaaaacagatttctaaaggaatacaaataggataagattgtgtatctgataat 731
QY 601 ttctctctgaaccagggttctgcatccactacacattgtcatgccaaattccacagaa 660
Db 732 ttctctctgaaccagggttctgcatccactacacattgtcatgccaaattccacagaa 791
QY 661 gctgtgagtccttcagtgctaccccttcagcttgcactgagcactgactgcttaataatgct 720
Db 792 gctgtgagtccttcagtgctaccccttcagcttgcactgagcactgactgcttaataatgct 851
QY 721 ataatgccttttagtaccttggaagaccttattcgatatcttgaaccagagatggcag 780
Db 852 ataatgccttttagtaccttggaagaccttattcgatatcttgaaccagagatggcag 911
QY 781 ttggactagaagatctatatagggcaacttggcaactcttggcaagcctttgtttt 840
Db 912 ttggactagaagatctatatagggcaacttggcaactcttggcaagcctttgtttt 971

Db 972 ggaagaaacccagagtggtgagctgaacctctaaacagaggagtggaattatcacgc 1031
QY 901 tgcacacctgtaacctctcagtgctcagaaggaagaaactaaagagacacgataccatt 960
Db 1032 tgcacacctgtaacctctcagtgctcagaaggaagaaactaaagagacacgataccatt 1091
QY 961 ttctggccaggttgctcctcgtggttaaacgctgtggtggaactgtgctgttgtctccac 1020
Db 1092 ttctggccaggttgctcctcgtggttaaacgctgtggtggaactgtgctgttgtctccac 1151
QY 1021 aattgcaatgaatgtcaatgtgtcccaagcaaaagttaactaaaaataaccacgaggtcctt 1080
Db 1152 aattgcaatgaatgtcaatgtgtcccaagcaaaagttaactaaaaataaccacgaggtcctt 1211
QY 1081 cagtgtagacccaagacgcgtgtcaggggattgcacaaatcactcacccagctgagccctg 1140
Db 1212 cagtgtagacccaagacgcgtgtcaggggattgcacaaatcactcacccagctgagccctg 1271
QY 1141 gacacacatgagagtgtagctgtgtgcagagggagacagagagtagccgcacac 1200
Db 1272 gacacacatgagagtgtagctgtgtgcagagggagacagagagtagccgcacac 1331
QY 1201 caccagcagctctgccacagagctgtgcagtllylgaagtggtgattctattagagaac 1260
Db 1332 caccagcagctctgccacagagctgtgcagt- ---gcagtggtgattctattagagaac 1387
QY 1261 gtatggttatcccatccttaactcagttgtgttgcctcagaagacctttcatcttcagg 1320
Db 1388 gtatggttatcccatccttaactcagttgtgttgcctcagaagacctttcatcttcagg 1447
QY 1321 attacagtgattctgaagagagagacatcaaacagaaattagagttgtgcaacagctc 1380
Db 1448 attacagtgattctgaagagagagacatcaaacagaaattagagttgtgcaacagctc 1507
QY 1381 ttttgagagaggcctaaagacagaggagaaaggcttccaatcgtggaagaaattataa 1440
Db 1508 ttttgagagaggcctaaagacagaggagaaaggcttccaatcgtggaagaaattataa 1567
QY 1441 tgtgtattataatagataccacagtagtttccagagttaccatgtacgtattccactagct 1500
Db 1568 tgtgtattataatagataccacagtagtttccagagttaccatgtacgtattccactagct 1627
QY 1501 ggtctgtatttcagttcttcgatacgcgttagggttaattgtagtgcagtagcaagaaac 1560
Db 1628 ggtctgtatttcagttcttcgatacgcgttagggttaattgtagtgcagtagcaagaaac 1687
QY 1561 tgtgcaagtgcacacctgattccgttgcctgttaactctaaagctccatgtcctgggc 1620
Db 1688 tgtgcaagtgcacacctgattccgttgcctgttaactctaaagctccatgtcctgggc 1747
QY 1621 ctataatcgataaaactggtgatt 1680
Db 1748 ctataatcgataaaactgga-ttt 1806
QY 1681 accagaactctctatgtactacaacacctggttttaaaaggaactatgttgcattgaat 1740
Db 1807 accagaactctctatgtactacaacacctggttttaaaaggaactatgttgcattgaat 1866
QY 1741 taaactgtgtcgtgctgatagga 1764
Db 1867 taaactgtgtcgtgctgatagga 1890

RESULT 12
AAC90564
ID AAC90564 standard; cDNA; 2849 BP.
XX
AC AAC90564;
XX
DT 21-MAR-2001 (first entry)
XX
DE Human PRO200 cDNA.

XX Human; PRO; cardiac; antiangiogenic; antiarteriosclerotic; hypotensive;
KW vasotropic; antirheumatic; antiarthritic; antiinflammatory; cytostatic;
KW vulnery; antianginal; gene therapy; cardiovascular disease;
KW endothelial disorder; angiogenic disorder; cancer; periodontal disease;
wound healing; ss.
XX Homo sapiens.
OS
XX WO200073445-A2.
PN
XX
PD 07-DEC-2000.
XX
PF 17-MAY-2000; 2000WO-US13705.
XX
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 26-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 01-SEP-1999; 99WO-US20111.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Gertschen ME;
PI Goddard A, Godowski PJ, Gurney AL, Kuo SS, Mark MR, Marsters SA;
PI Paoni NF, Pitti RM, Watanabe CK, Williams PM, Wood WI;
XX
DR WPI: 2001-025251/03.
DR P-PSDB; AAB50952.
XX
XX Seventeen nucleic acids encoding PRO polypeptides which are useful in
PT diagnosis and treatment of cardiovascular, endothelial or angiogenic
PT disorders in a mammal -
XX
PS Claim 60; Fig 3; 182pp; English.
XX
CC The present sequence is one of seventeen nucleic acids encoding PRO
CC polypeptides. The PRO nucleic acids, polypeptides, agonists and
CC antagonists are useful for treating cardiovascular, endothelial or
CC angiogenic disorders in a mammal. Examples of these disorders include
CC cardiac hypertrophy, trauma, cancer, age-related macular degeneration,
CC atherosclerosis, hypertension, arterial restenosis, Reynaud's disease,
CC rheumatoid arthritis, angina, myocardial infarctions, thrombophlebitis
CC and lymphangitis. The PRO polypeptides and antagonists are also used to
CC prevent tumour angiogenesis and for treating periodontal diseases. They
CC are also used to stimulate wound healing and tissue regeneration.
CC The PRO nucleic acids, polypeptides and anti-PRO antibodies are useful
CC for diagnosing a cardiovascular, endothelial or angiogenic disorder.
XX
SQ Sequence 2849 BP; 851 A; 528 C; 619 G; 850 T; 1 other;

Query Match 98.2%; Score 1732.4; DB 22; Length 2849;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;
QY 1 attatgtggaactaccctgcattctctgctgcagagcagctgcgctccacccc 60
XX
Db 132 attatgtggaactaccctgcattctctgctgcagagcagctgcgctccacccc 191

| | | | |
|----|-----|--|-----|
| Qy | 661 | gctgtgagttccttoagtgtaccccccttcagcttttggccactggacctgttaataatgct | 720 |
| Db | 662 | | 721 |
| Db | 663 | | 722 |
| Db | 664 | | 723 |
| Db | 665 | | 724 |
| Qy | 666 | gctgtgagttccttcagtgctacccccctcagctttgccactggacctgttaataatgct | 725 |
| Qy | 667 | | 726 |
| Qy | 668 | | 727 |
| Qy | 669 | | 728 |
| Qy | 670 | ataactgccttttagtaaccttggagacctattcgatatcttgaaccagagagatggcag | 729 |
| Db | 671 | | 730 |
| Db | 672 | | 731 |
| Db | 673 | | 732 |
| Db | 674 | ataactgccttttagtaaccttggagacctattcgatatcttgaaccagagagatggcag | 733 |
| Qy | 675 | | 734 |
| Qy | 676 | | 735 |
| Qy | 677 | | 736 |
| Qy | 678 | | 737 |
| Qy | 679 | | 738 |
| Qy | 680 | | 739 |
| Qy | 681 | tttgacttagaagatctatataggccaaacttggcaacttcttggcaaggctttgttttt | 740 |
| Db | 682 | | 741 |
| Db | 683 | | 742 |
| Db | 684 | | 743 |
| Db | 685 | tttgacttagaagatctatataggccaaacttggcaacttcttggcaaggctttgttttt | 744 |
| Qy | 686 | | 745 |
| Qy | 687 | | 746 |
| Qy | 688 | | 747 |
| Qy | 689 | | 748 |
| Qy | 690 | ggaagaaaaatccagagtggttgagatctgaacctcttaacagagaggttaagattacagc | 749 |
| Db | 691 | | 750 |
| Db | 692 | ggaagaaaaatccagagtggttgagatctgaacctcttaacagagaggttaagattacagc | 751 |
| Qy | 693 | | 752 |
| Qy | 694 | | 753 |
| Qy | 695 | | 754 |
| Qy | 696 | | 755 |
| Qy | 697 | | 756 |
| Qy | 698 | | 757 |
| Qy | 699 | | 758 |
| Qy | 700 | | 759 |
| Qy | 701 | tgcaacctcgtaacctctcagtgctcataagggaagaataaaagaaacgataccatt | 760 |
| Db | 702 | | 761 |
| Db | 703 | | 762 |
| Db | 704 | tgcaacctcgtaacctctcagtgctcataagggaagaataaaagaaacgataccatt | 763 |
| Qy | 705 | | 764 |
| Qy | 706 | | 765 |
| Qy | 707 | | 766 |
| Qy | 708 | | 767 |
| Qy | 709 | | 768 |
| Qy | 710 | | 769 |
| Qy | 711 | | 770 |
| Qy | 712 | | 771 |
| Qy | 713 | | 772 |
| Qy | 714 | | 773 |
| Qy | 715 | | 774 |
| Qy | 716 | | 775 |
| Qy | 717 | | 776 |
| Qy | 718 | | 777 |
| Qy | 719 | | 778 |
| Qy | 720 | | 779 |
| Qy | 721 | | 780 |
| Qy | 722 | | 781 |
| Qy | 723 | | 782 |
| Qy | 724 | | 783 |
| Qy | 725 | | 784 |
| Qy | 726 | | 785 |
| Qy | 727 | | 786 |
| Qy | 728 | | 787 |
| Qy | 729 | | 788 |
| Qy | 730 | | 789 |
| Qy | 731 | | 790 |
| Qy | 732 | | 791 |
| Qy | 733 | | 792 |
| Qy | 734 | | 793 |
| Qy | 735 | | 794 |
| Qy | 736 | | 795 |
| Qy | 737 | | 796 |
| Qy | 738 | | 797 |
| Qy | 739 | | 798 |
| Qy | 740 | | 799 |
| Qy | 741 | | 800 |
| Qy | 742 | | 801 |
| Qy | 743 | | 802 |
| Qy | 744 | | 803 |
| Qy | 745 | | 804 |
| Qy | 746 | | 805 |
| Qy | 747 | | 806 |
| Qy | 748 | | 807 |
| Qy | 749 | | 808 |
| Qy | 750 | | 809 |
| Qy | 751 | | 810 |
| Qy | 752 | | 811 |
| Qy | 753 | | 812 |
| Qy | 754 | | 813 |
| Qy | 755 | | 814 |
| Qy | 756 | | 815 |
| Qy | 757 | | 816 |
| Qy | 758 | | 817 |
| Qy | 759 | | 818 |
| Qy | 760 | | 819 |
| Qy | 761 | | 820 |
| Qy | 762 | | 821 |
| Qy | 763 | | 822 |
| Qy | 764 | | 823 |
| Qy | 765 | | 824 |
| Qy | 766 | | 825 |
| Qy | 767 | | 826 |
| Qy | 768 | | 827 |
| Qy | 769 | | 828 |
| Qy | 770 | | 829 |
| Qy | 771 | | 830 |
| Qy | 772 | | 831 |
| Qy | 773 | | 832 |
| Qy | 774 | | |

Db 1867 taaactgtctatctgtatagga 1890
 |||||

RESULT 14

AAC97404

ID AAC97404 standard; cDNA; 2849 BP.

AC AAC97404;

DT 28-FEB-2001 (first entry)

DE Human angiogenesis-associated protein PRO200 cDNA, SEQ ID NO:50.

XX Human; angiogenesis-associated protein; PRO; endothelial cell growth;
 KW cardiac hypertrophy; cardiovascular disorder; endothelial disorder;
 KW angiogenic disorder; atherosclerosis; osteoporosis; hypertension;
 KW myocardial infarction; diabetic retinopathy; rheumatoid arthritis;
 KW Crohn's disease; psoriasis; endometriosis; ulcer; wound healing; cancer;
 KW Alzheimer's disease; Huntington's disease; stroke; drug screening;
 KW gene therapy; transgenic animal; ss.

OS Homo sapiens.

XX WO2000053753-A2.

XX 14-SEP-2000.

XX 05-JAN-2000; 2000WO-US00219.

XX 08-MAR-1999; 99WO-US05028.

XX 12-MAR-1999; 99US-0123957.

XX 14-MAY-1999; 99US-0134287.

XX 02-JUN-1999; 99WO-US12252.

XX 23-JUN-1999; 99US-0141037.

XX 20-JUL-1999; 99US-0144756.

XX 26-JUL-1999; 99US-0145698.

XX 01-SEP-1999; 99WO-US20111.

XX 08-SEP-1999; 99WO-US20594.

XX 15-SEP-1999; 99WO-US21090.

XX 15-SEP-1999; 99WO-US21547.

XX 05-OCT-1999; 99WO-US23089.

XX 30-NOV-1999; 99WO-US28313.

XX 30-NOV-1999; 99WO-US28409.

XX 02-DEC-1999; 99WO-US28564.

XX 02-DEC-1999; 99WO-US28565.

XX (GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Goddard A;

PI Godowski PJ, Gurney AL, Hillan KJ, Kuo SS, Mark MR, Marsters SA;

PI Paoni NF, Pitti RM, Watanabe CK, Williams PM, Wood WI;

XX WPI; 2001-090793/10.

XX P-PSDB; AAB53074.

XX New isolated nucleic acid for producing a PRO polypeptide, analyzing
 PT genetic disorders and treating cardiovascular, endothelial or
 PT angiogenic disorders, such as atherosclerosis, wounds or cancer -

XX Claim 58; Fig 21A-B; 293pp; English.

XX The invention relates to novel human angiogenesis-associated proteins
 CC designated PRO proteins (AAB53064-B53097), and to nucleic acids encoding
 CC PRO proteins. The invention also relates to vectors and host cells
 CC comprising a PRO nucleic acid, the recombinant production of a PRO
 CC protein, PRO antibodies specific for a PRO protein, fusion proteins
 CC comprising a PRO protein, agonists or antagonists of a PRO protein, and
 CC compounds which inhibit the expression of a PRO gene. The invention
 CC additionally encompasses methods of identifying modulators of PRO
 CC expression or activity; diagnosing a cardiovascular, endothelial or
 CC angiogenic disorder, or a susceptibility to such a disorder by detecting
 CC mutations in a PRO gene, or the expression level of a PRO gene within a

CC particular tissue; treating a cardiovascular, endothelial or angiogenic
 CC disorder via the administration of a PRO protein, PRO nucleic acid, or
 CC PRO agonist or antagonist; a retroviral gene therapy vector comprising a
 CC PRO nucleic acid; and methods of inhibiting or stimulating endothelial
 CC cell growth, cardiac hypertrophy or PRO-induced angiogenesis via the
 CC administration of a PRO protein, or an agonist or antagonist thereof.
 CC PRO nucleic acids, PRO proteins, or antibodies against PRO proteins, PRO
 CC agonists and PRO antagonists may be used as therapeutic agents to treat
 CC cardiovascular, endothelial or angiogenic disorders, such as
 CC atherosclerosis, osteoporosis, myocardial infarction, hypertension,
 CC diabetic retinopathy, rheumatoid arthritis, Crohn's disease, psoriasis,
 CC endometriosis, ulcers, wounds, cancer, Alzheimer's disease, Huntington's
 CC disease, or stroke. PRO nucleic acids are additionally useful in the
 CC recombinant production of PRO proteins, as hybridisation probes to
 CC screen libraries to isolate cDNAs with sequence identity to PRO proteins,
 CC to map genes encoding PRO proteins, to analyse genetic disorders, and in
 CC gene therapy. PRO nucleic acids can also be used to produce transgenic
 CC animals useful for the development and screening of potential
 CC therapeutic agents. The present sequence represents a cDNA encoding a PRO
 CC protein of the invention.
 XX
 SQ Sequence 2849 BP; 851 A; 528 C; 619 G; 850 T; 1 other;

Query Match 98.2%; Score 1732.4; DB 22; Length 2849;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1758; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

QY 1 attatgtgaaacacccctgctgattctctgtccagagcagcgtcgccctccacccc 60

|||||

Db 132 attatgtgaaacacccctgctgattctctgtccagagcagcgtcgccctccacccc 191

|||||

QY 61 agtcgagcctccctggcggtgtgaaagagactcgggagtcgtctccaaagtgc 120

|||||

Db 192 agtcgagcctccctggcggtgtgaaagagactcgggagtcgtctccaaagtgc 251

|||||

QY 121 cgcctgagtgagctctcaccacagtcagccaaatgagcctcttcgggctctcctgctg 180

|||||

Db 252 cgcctgagtgagctctcaccacagtcagccaaatgagcctcttcgggctctcctgctg 311

|||||

QY 181 acatctgacctggcggccagagacaggggactcaggggaatcccaacctgagtagtaaa 240

|||||

Db 312 acatctgacctggcggccagagacaggggactcaggggaatcccaacctgagtagtaaa 371

|||||

QY 241 ttccagttttccagcaacaaaggaacagcggagtcacaaatcctcagcagtagagaatt 300

|||||

Db 372 ttccagttttccagcaacaaaggaacagcggagtcacaaatcctcagcagtagagaatt 431

|||||

QY 301 attactgtgtctactaatggaagtattccagcccaaggtttcctcactatccaaga 360

|||||

Db 432 attactgtgtctactaatggaagtattccagcccaaggtttcctcactatccaaga 491

|||||

QY 361 aatacggcttggtgtagattagtcagtagagagaaaattatggatacaactaacg 420

|||||

Db 492 aatacggcttggtgtagattagtcagtagagagaaaattatggatacaactaacg 551

|||||

QY 421 ttgatgaagatttggtcttgagaccagagatgacatatgcaagtattgttga 480

|||||

Db 552 ttgatgaagatttggtcttgagaccagagatgacatatgcaagtattgttga 611

|||||

QY 481 gaagttgaggaacccagtgatggaactatattaggcgctggtggtgtctggtactgta 540

|||||

Db 612 gaagttgaggaacccagtgatggaactatattaggcgctggtggtgtctggtactgta 671

|||||

QY 541 ccaggaaaacagatttctaaagaaatcaaataggataagatttgtatctgtagtaata 600

|||||

Db 672 ccaggaaaacagatttctaaagaaatcaaataggataagatttgtatctgtagtaata 731

|||||

QY 601 ttctcttgaaacaggggtctgcacccactacaacattgtcatgcacaaattccagaa 660

|||||

Db 732 ttctcttgaaacaggggtctgcacccactacaacattgtcatgcacaaattccagaa 791

|||||

QY 661 gctgtgagctctcagtgctaccccttcagctttgccactggacctgcttaataatgct 720

3

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic-search, using sw model

Run on: June 30, 2001, 00:41:33 ; Search time 97.06 Seconds
(without alignments)
3366.982 Million cell updates/sec

Title: US-09-457-066-1

Perfect score: 1764

Sequence: 1 attatgtgaaactaccctg.....cttgtctgctgatagga 1764

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 317530 seqs, 92630164 residues

Total number of hits satisfying chosen parameters: 635060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| C 1 | 42.6 | 2.4 | 7218 | 1 | US-08-232-463-14 |
| 2 | 37.2 | 2.1 | 3098 | 1 | US-08-447-500-1 |
| 3 | 37.2 | 2.1 | 3098 | 1 | US-08-454-097-1 |
| 4 | 37.2 | 2.1 | 3098 | 1 | US-08-447-408-1 |
| 5 | 37.2 | 2.1 | 3098 | 1 | US-08-453-866-1 |
| 6 | 37.2 | 2.1 | 3098 | 3 | US-08-185-359-1 |
| C 7 | 35.6 | 2.0 | 3581 | 2 | US-08-738-349-1 |
| C 8 | 35.2 | 2.0 | 29604 | 3 | US-08-781-891-207 |
| C 9 | 35 | 2.0 | 1931 | 3 | US-09-019-942-2 |
| C 10 | 34 | 1.9 | 1959 | 4 | US-09-201-641-1 |
| 11 | 33.8 | 1.9 | 10684 | 3 | US-08-618-100B-3 |
| 12 | 33.6 | 1.9 | 1992 | 1 | US-08-455-550-6 |
| C 13 | 33.6 | 1.9 | 2798 | 2 | US-08-870-518-6 |
| 14 | 33.2 | 1.9 | 528 | 3 | US-08-714-918-24 |
| 15 | 33.2 | 1.9 | 528 | 4 | US-09-265-315-24 |
| 16 | 33.2 | 1.9 | 528 | 4 | US-09-265-315-24 |
| 17 | 33.2 | 1.9 | 528 | 4 | US-09-286-417-24 |
| C 18 | 33.2 | 1.9 | 1841 | 5 | PCT-US95-00362-1 |
| C 19 | 33 | 1.9 | 1875 | 5 | PCT-US96-10618-1 |
| C 20 | 32.8 | 1.9 | 4046 | 1 | US-07-793-961A-1 |
| C 21 | 32.8 | 1.9 | 4046 | 1 | US-08-240-357-1 |
| C 22 | 32.8 | 1.9 | 56516 | 2 | US-08-996-306-1 |
| 23 | 32.6 | 1.8 | 519 | 3 | US-08-714-918-67 |
| 24 | 32.6 | 1.8 | 519 | 4 | US-09-265-315-67 |
| 25 | 32.6 | 1.8 | 519 | 4 | US-09-265-315-67 |
| 26 | 32.6 | 1.8 | 519 | 4 | US-09-266-417-67 |
| 27 | 32.4 | 1.8 | 7218 | 1 | US-08-232-463-14 |

| | | | | | | |
|------|------|-----|------|---|-------------------|-------------------|
| 28 | 32.2 | 1.8 | 7244 | 4 | US-08-378-313-26 | Sequence 26, Appl |
| C 29 | 32 | 1.8 | 2007 | 3 | US-08-747-221B-36 | Sequence 36, Appl |
| 30 | 32 | 1.8 | 2007 | 3 | US-08-747-221B-38 | Sequence 38, Appl |
| 31 | 31.8 | 1.8 | 1519 | 1 | US-07-971-759-19 | Sequence 19, Appl |
| C 32 | 31.8 | 1.8 | 2555 | 2 | US-08-693-457-3 | Sequence 3, Appl |
| C 33 | 31.8 | 1.8 | 2555 | 4 | US-09-265-731-3 | Sequence 3, Appl |
| C 34 | 31.8 | 1.8 | 5852 | 3 | US-07-867-106-2 | Sequence 2, Appl |
| C 35 | 31.6 | 1.8 | 857 | 3 | US-08-460-040-1 | Sequence 1, Appl |
| C 36 | 31.6 | 1.8 | 1172 | 1 | US-07-945-288-9 | Sequence 1, Appl |
| C 37 | 31.6 | 1.8 | 1172 | 1 | US-08-462-831-9 | Sequence 9, Appl |
| C 38 | 31.6 | 1.8 | 1172 | 1 | US-08-461-809-9 | Sequence 9, Appl |
| C 39 | 31.6 | 1.8 | 1172 | 1 | US-08-461-441-9 | Sequence 9, Appl |
| C 40 | 31.6 | 1.8 | 1172 | 5 | PCT-US93-08518-9 | Sequence 9, Appl |
| C 41 | 31.6 | 1.8 | 1524 | 1 | US-08-197-792-34 | Sequence 34, Appl |
| C 42 | 31.6 | 1.8 | 1524 | 1 | US-08-459-850-34 | Sequence 34, Appl |
| C 43 | 31.6 | 1.8 | 1524 | 1 | US-08-459-214-34 | Sequence 34, Appl |
| C 44 | 31.6 | 1.8 | 1797 | 3 | US-08-946-026-13 | Sequence 13, Appl |
| 45 | 31.2 | 1.8 | 1411 | 4 | US-09-387-574-1 | Sequence 1, Appl |

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; CLONE: ptzgpt-F1s
US-08-232-463-14

IMMEDIATE SOURCE:
CLONE: Protein Kinase
FEATURE:
NAME/KEY: CDS
LOCATION: 879..2360
US-08-454-097-1

Query Match 2.1%; Score 37.2; DB 1; Length 3098;

Best Local Similarity 54.3%; Pred. No. 0.59; Mismatches 0; Gaps 0;

Matches 75; Conservative 0; Indels 63; Length 3098;

QY 1296 gcttcaaggacattcttcattcaggtattacagtgattctgaaagagagagacatcaaac 1355

Db 572 GTTTGAATTTCTTTTCTGGAGAAATTTTGGTGCACGAGGAGGAGAGAGAGAA 631

QY 1356 agaattaggagttgtgaacagctcttttgagagagagcctaaagagagagagagaggt 1415

Db 632 AAAAGTTGAACACGACCATATATGACGTGGTTGAAATACAAAGAGAGAGAAAGGTT 691

QY 1416 cttcaatcgtgaaagaa 1433

Db 692 CGACACTCGAGGAAAGCA 709

RESULT 4

US-08-447-408-1

Sequence 1, Application US/08447408

Patent No. 5705377

GENERAL INFORMATION:

APPLICANT: Hoekstra, Merl F.

TITLE OF INVENTION: TYROSINE KINASE

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Spensley Horn Jubas & Lubitz

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/447,408

FILING DATE: 03-JUL-1991

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/728,783

FILING DATE: 03-JUL-1991

ATTORNEY/AGENT INFORMATION:

NAME: Wetherell Ph.D., John R.

REGISTRATION NUMBER: 31,678

REFERENCE/DOCKET NUMBER: PD-1318

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 455-5100

TELEFAX: (619) 455-5110

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3098 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

IMMEDIATE SOURCE:

CLONE: Tyrosine Kinase

FEATURE:

NAME/KEY: CDS

LOCATION: 879..2364

US-08-447-408-1

Query Match 2.1%; Score 37.2; DB 1; Length 3098;

Best Local Similarity 54.3%; Pred. No. 0.59; Mismatches 0; Gaps 0;

Matches 75; Conservative 0; Indels 63; Length 3098;

QY 1296 gcttcaaggacattcttcattcaggtattacagtgattctgaaagagagagacatcaaac 1355

Db 572 GTTTGAATTTCTTTTCTGGAGAAATTTTGGTGCACGAGGAGGAGAGAGAGAA 631

QY 1356 agaattaggagttgtgaacagctcttttgagagagagcctaaagagagagagagaggt 1415

Db 632 AAAAGTTGAACACGACCATATATGACGTGGTTGAAATACAAAGAGAGAGAAAGGTT 691

QY 1416 cttcaatcgtgaaagaa 1433

Db 692 CGACACTCGAGGAAAGCA 709

RESULT 5

US-08-453-866-1

Sequence 1, Application US/08453866

Patent No. 5756289

GENERAL INFORMATION:

APPLICANT: Hoekstra, Merl F.

TITLE OF INVENTION: PROTEIN KINASES

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Spensley Horn Jubas & Lubitz

STREET: 1880 Century Park East, Suite 500

CITY: Los Angeles

STATE: CA

COUNTRY: USA

ZIP: 90067

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/453,866

FILING DATE: 30-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/008,001

FILING DATE: 20-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Wetherell Ph.D., John R.

REGISTRATION NUMBER: 31,678

REFERENCE/DOCKET NUMBER: PD-2458

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 455-5100

TELEFAX: (619) 455-5110

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3098 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

IMMEDIATE SOURCE:

CLONE: Protein Kinase

FEATURE:

NAME/KEY: CDS

LOCATION: 879..2360

US-08-453-866-1

Query Match

Best Local Similarity 54.3%; Score 37.2; DB 1; Length 3098;

Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1296 gcttcaaggacattcttcattcaggtattacagtgattctgaaagagagagacatcaaac 1355

Db 572 GTTTGAATTTCTTTTCTGGAGAAATTTTGGTGCACGAGGAGGAGAGAGAGAA 631

QY 1356 agaattaggagttgtgaacagctcttttgagagagagcctaaagagagagagagaggt 1415

Db 632 AAAAGTTGAACACGACCATATATGACGTGGTTGAAATACAAAGAGAGAGAAAGGTT 691

QY 1416 cttcaatcgtgaaagaa 1433

Db 692 CGACACTCGAGGAAAGCA 709

Db 632 AAAAGTTGAACACGACCCACATATATGGAACGTGTTGAAATACAAGAGAAGAAGGTT 691

Qy 1416 cttcaatcgtaggaagaa 1433

Db 692 CGACACTCGAGGAAAGCA 709

RESULT⁷
US-08-738-349-1/C
; Sequence 1, Application US/08738349
; Patent No. 5869638
; GENERAL INFORMATION:
; APPLICANT: Takeshita, Sunao
; APPLICANT: Okazaki, Makoto
; APPLICANT: Kawai, Shinji
; APPLICANT: Tsujimura, Atsushi
; APPLICANT: Amann, Egon
; TITLE OF INVENTION: Bone-Related Cadherin-Like Protein and
; TITLE OF INVENTION: Process for Its Production
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.

COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: DS/08/738,349

APPLICATION NUMBER: US/08/138,343
FILING DATE: 25-OCT-1996

CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 03 08/304,433
FILING DATE:

AFFLICTION NUMBER: 03 06/112,001
FILING DATE: 26-AUG-1993

ATTORNEY/AGENT INFORMATION.
NAME: Barker, M P

REGISTRATION NUMBER: 32,013
REFERENCE/DOCKET NUMBER: 02481 13

TELEPHONE: 202-408-4000

TELEFAX: 202 408 4400
INFORMATION FOR SEO ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 3581 base pairs

TYPE: nucleic acid
STRANDEDNESS: single

MOLECULE TYPE: cdna to mRNA

ORIGINAL SOURCE: MUS MUSCULUS

SINAIN: OSTEOPLASTIC CELL LINE MODEL
FEATURE:

| | |
|-----------|----------|
| NAME/KEY: | CDS |
| LOCATION: | 284 2671 |

4
3
3
3
3
3

2.08: score 3

Local Similarity 04:00, 113d: 1
 Mismatches 53: Conservative 0: Mismatches

1643 ttttttttttttttttttqctcatattcaca

3539

Search completed: June 30, 2001, 01:49:04
Job time: 4051 sec

US-09-265-315-24
 : Sequence 24, Application US/09265315
 : Patent No. 6187541
 : GENERAL INFORMATION:
 : APPLICANT: Benton, Bret
 : APPLICANT: Lee, Ving J.
 : APPLICANT: Malouin, Francois
 : APPLICANT: Martin, Patrick K.
 : APPLICANT: Schmid, Molly B.
 : APPLICANT: Sun, Dongxu
 : TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
 : TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
 : NUMBER OF SEQUENCES: 111
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Lyon & Lyon
 : STREET: 633 West Fifth Street
 : STREET: Suite 4700
 : CITY: Los Angeles
 : STATE: California
 : COUNTRY: U.S.A.
 : ZIP: 90071-2066
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 : MEDIUM TYPE: storage
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: IBM P.C. DOS 5.0
 : SOFTWARE: Word Perfect 5.1
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/265,315
 : FILING DATE: March 9, 1999
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/714,918
 : FILING DATE: September 13, 1996
 : APPLICATION NUMBER: 60/009,102
 : FILING DATE: December 22, 1995
 : APPLICATION NUMBER: 60/003,798
 : FILING DATE: September 15, 1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Warburg, Richard J.
 : REGISTRATION NUMBER: 32,327
 : REFERENCE/DOCKET NUMBER: 240/247
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (213) 489-1600
 : TELEFAX: (213) 955-0440
 : TELEX: 67-3510
 : INFORMATION FOR SEQ ID NO: 24:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 528 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : US-09-265-315-24

| | Query Match | 1.9%; | Score 33.2; | DB 4; | Length 528; |
|-----|-----------------------|---|----------------|------------|-------------|
| | Best Local Similarity | 50.0%; | Pred. No. 3.3; | Mismatches | Indels |
| | Matches 71; | Conservative 0; | Mismatches 71; | Indels 0; | Gaps 0; |
| QY | 300 | tattactgtctactaaatgaagtattcacagcccaaggtttcttcatacattatccaag | 359 | | |
| Ddb | 343 | TAAAGTTTGCCCTTATTTCACCAATACATTGGAATNGNCITCGTNAGNCAATTGTAAAG | 402 | | |
| QY | 360 | aatatcggtcttggttagagattagtagcagtagagaaaaatgtaggcatacaacttac | 419 | | |
| Ddb | 403 | AGATNCGGGCATAAATTTTGTCGCCACAATCAATTTGGTATTCTTCTGTACGCCTAC | 462 | | |
| QY | 420 | gtttgatgaagaattgggctt | 441 | | |
| Dd | 463 | GGTNAATTAAATACCNTGGNTT | 484 | | |

3